Royal Netherlands
Meteorological Institute
*Ministry of Infrastructure
and Water Management*

Advice on the computation of peak-ground-velocity confidence regions for events in gas fields other than the Groningen gas field

E. Ruigrok and B. Dost

Advice on the computation of peak-ground-velocity confidence regions for events in gas fields other than the Groningen gas field

KNMI, R&D Seismology and Acoustics

June 30, 2020

Summary

On request by the Technical committee on ground movement (Tcbb), a study was carried out to define the region affected by induced earthquakes related to gas exploitation[1] from fields other than the Groningen gas field. This request is based on advise that the Tcbb formulated in 2019, on a national approach for handling damage due to mining.

We built a database of a few thousand peak-ground velocity (PGV) recordings, both from induced events inside and outside Groningen. Subsequently, we assessed which existing ground-motion model best fits this Dutch PGV database.

It was found that the empirical PGV model for Groningen (*Bommer et al.*, 2019) was equally suited for induced events in- and outside Groningen and showed the best fit to the database. There was no marked difference for events with and without a shielding overburden. The selected model was adapted at short distance and for low magnitudes, outside the applicable magnitude range of the original model, to obtain an improved fit to the PGV database. A depth dependence was added to make the updated model (BMR2) applicable to source regions with nucleation depths other than 3 km. The BMR2 model has the same aleatory variability as the original empirical PGV model. This variability incorporates, among others, epicentral uncertainty and variation in site effects, radiation pattern and stress drop.

The BMR2 model, together with local recordings, is used for computing PGV contours for 2 mm/s until the maximum level, with a 1, 10 and 50% probability of exceedance. In the current PGV database, the lowest magnitude event for which 2 mm/s has been measured, is 1.9. For a 1% probability of exceedance, the BMR2 model reaches 2 mm/s at a local magnitude M of 1.53. Therefore, only for $M \geq 1.5$ or when a PGV recording is larger than 2 mm/s, a PGV threshold region computation is started up. The recordings are used to estimate and remove an event term. Aside from a small nugget, the local PGV recordings are used as ground truths at their recording site and away from that site they are combined with the model to obtain a combined PGV field. The radius of influence of these recordings, that is the distance at which the weight of the model and recording is equal, is computed with the PGV database to be 2.7 km.

The computation of PGV contours is illustrated numerically and with three events that occurred outside the Groningen gas field. For one of these events (Dalen 17-7-2018) the event term could largely be computed, but there were no near-source recordings to locally perturb the modeled PGV field. For the Warder 04-06-2018 event there was a lack of both local and regional PGV recordings which made it necessary to use the full variability of the model, leading to relatively large PGV threshold regions. For the Roswinkel 25-10-2000 event, the availability of 4 recordings within a few kilometers from the source led to a near-epicentre PGV field dominated by the measurements.

---

[1]Although the request was specifically for induced seismicity related to gas exploitation, the approach outlined in this report would also be valid for induced seismicity related to other anthropogenic perturbations in the upper crust

1 Introduction

Earthquakes in the Netherlands are monitored with geophones, broadbands and accelerometers. The accelerometers are placed on the Earth's surface and serve to quantify the ground motion that could inflict damage. They are referred to as strong-motion instruments since their dynamic range has been set to enable recording up till very large accelerations, exceeding the gravity acceleration. The strong-motion network in the Netherlands is well developed in the northeast of the Netherlands. This has led to a rich database of ground-motion recordings which is used to develop models which are tailored to induced seismicity in Dutch subsurface settings. This network is continuously expanding to better cover regions outside of Groningen, to better quantify ground motion due to induced events in other parts of the country. Data from this expanding network will play a key role in the procedures proposed in this report and are referred to as "local" data.

KNMI was asked by the Technical committee on ground movement (Technische commissie bodem beweging, Tcbb) to develop a workflow for computing ground-motion contours for induced seismic events in gas fields other than the Groningen gas field. Specifically the request refers to the horizontal Peak Ground Velocity (PGV) values for a 1%, 10% and 50% probability of exceedance for PGV levels of 2, 3, 4, 5, 10 mm/s until the maximum level, with steps of 5 mm/s and for different definitions of PGV.

Structures are mostly sensitive to seismic waves in the frequency range in which these structures resonate. The PGV is a proxy for the ground motion in this frequency range and is widely used to assess potential earthquake damage. Also the Dutch guideline on assessing building damage (*Jonker et al.*, 2017) is expressed in terms of PGV. The vibrations are typically recorded in three orthogonal directions. In the Dutch guideline, PGVmax is used, which is the maximum absolute value in one of the three recorded directions. For buildings, the guideline lists PGV threshold values that yield a 1% probability of inflicting damage. These PGV values range from 3 mm/s for sensitive brickwork, to 5 mm/s for conventional brickwork and to 20 mm/s for concrete and wooden structures. For repeated occurrence, lower PGVs could inflict damage (Table 10.6 in *Jonker et al.* (2017)).

For events in Groningen, a detailed ground-motion model has been derived (*Bommer et al.*, 2017) and a dense network of surface recordings is available (*Dost et al.*, 2017). This allows computation of so-called shakemaps (*Wald et al.*, 2006) which show the ground motion at the Earth's surface: `https://www.knmi.nl/nederland-nu/seismologie/aardbevingen/shakemaps-archief`. These can be used, among others, for damage assessment. In addition, a simplified version of the ground-motion model for Groningen was developed, namely a fully empirical PGV model, only valid for a limited magnitude range (M 1.8-3.6) (*Bommer et al.*, 2019). This model is a function of distance and magnitude only.

For events outside Groningen, the measuring network is more sparse and site conditions are not well known. In this case, the local recordings that are present are descriptive of local PGV levels, but for large parts of the region one needs to resort to generic ground-motion-

prediction equation (GMPE) with a variability, that incorporates the large differences in, among others, site conditions.

A GMPE is a function that describes the decay in amplitude with distance and scaling with magnitude. Sometimes more variables are included, like site class, style-of-faulting and focal depth. A large suite of GMPEs has been derived for various settings. Three empirical GMPEs exist that have been derived for seismicity with low magnitudes (<5) and shallow focal depths (<10 km) and that have been calibrated, purely or partly, with Dutch data. The first such model is described in *Dost et al.* (2004) in which seismicity outside Groningen is used for calibration. The second model (*Douglas et al.*, 2013) uses partly the same Dutch data as in *Dost et al.* (2004), but adds to this data from geothermal areas in Europe and USA. The third model has been calibrated with data from Groningen only (*Bommer et al.*, 2019). We consider a fourth model, the one from *Atkinson* (2015), which has been derived for induced seismicity, using a database of shallow tectonic events. These shallow events are generally deeper than induced events and there are little recordings at short ranges. The models are averages over many different sites and earthquakes. These variations add to the variability in the models, which has been quantified in all four of them.

In the following we derive a database of a few thousand PGVs from past recordings. We use this database to assess which GMPE best describes the attenuation observed in the Dutch subsurface setting. We do this assessment separately for events related to the Groningen gas field and all other induced events. The best GMPE is then adapted at short range to give an improved match with the PGV database. The adaptation is inspired by the radiation pattern of a typical normal-faulting event at 3 km depth. This adapted model forms the base for computation of PGV threshold regions. Thereafter, we discuss the combination of local recordings and a GMPE for computing PGV threshold regions. The distance at which local recordings are weighted in, is controlled by a distance-standard-deviation model which is obtained from the PGV database. Next, we outline a general workflow and exemplify it with an application to three events that occurred outside of Groningen.

This report contains some redundancy with an earlier report (*Ruigrok and Dost*, 2020) in which a workflow was developed for one specific event outside Groningen: the Zuidlaren 23-12-2016 event. This report contains a more generic analysis and includes the treatment of PGV recordings that are near the epicenter that locally perturb the model-based contours.

2 Database

In this chapter we describe the data on which the analysis in the following chapters is based. The source of the PGV data is the recording network, which is shortly described in the first section. The settings with which the PGVs are obtained are described in detail in the second section to facilitate independent checks. The third section shows the spatial PGV distribution due to the radiation pattern of a characteristic induced event at 3 km depth. Insights from this modelling are used when making adaptions to an existing GMPE.

2.1 Network

Until the end of the nineteen eighties, seismicity was only known to exist in the southern part of the Netherlands. This part of the country is located within the lower Rhine Graben, which is an extensional setting characterized by normal faulting. As a consequence, there was a seismic network in the south of the country and only one broadband station in the upper half of the country (Witteveen, Drenthe).

The first event in the north of the Netherlands was detected in 1986 in Assen. This event prompted the installation of (temporary) seismic networks. The Assen event was followed by other events near gas fields, in North Holland, Drenthe and Groningen. In 1993 these events were concluded to be induced (*Haak*, 1993).

Halfway through the nineties, a fixed geophone network was built in the northeast of the country, specifically to monitor induced seismicity. Also an accelerometer network was laid out, to record ground motion. This network of GeoSig accelerometers existed between 1997 to halfway 2013 and operated in a triggered mode. Since September 2013, EpiSensor accelerometers are in operation that record in a continuous fashion. Details on this network in the Groningen area can be found in *Dost et al.* (2017). All data are publicly available.

Figure 2.1 shows all stations that are currently (Feb. 2020) in operation and whose data are directly accessible through the KNMI website. Most of these sensors are operated by the KNMI. A small part of the sensors is operated by other parties, but the data are directly streamed to the KNMI. The total number of accelerometers is 123. This strong-motion network will be further expanded in 2020, with around 30 new accelerometers. The coverage will be improved in areas outside Groningen where gas is produced or where geothermal fields are in operation: Zuid-Holland, Noord-Holland, Friesland, Drenthe, Overijssel and Brabant. Moreover, a strong-motion network will be built in East Brabant and Limburg with a focus on quantifying ground motion due to tectonic events.

There are also strong-motion sensors of which the data are not directly accessible by KNMI. Different industries have installed networks at their sites for their own use. Moreover, in the Groningen area there is a densely spaced household network operated by TNO. It contains more than 300 accelerometers. Details on all ground-motion networks in Groningen can be
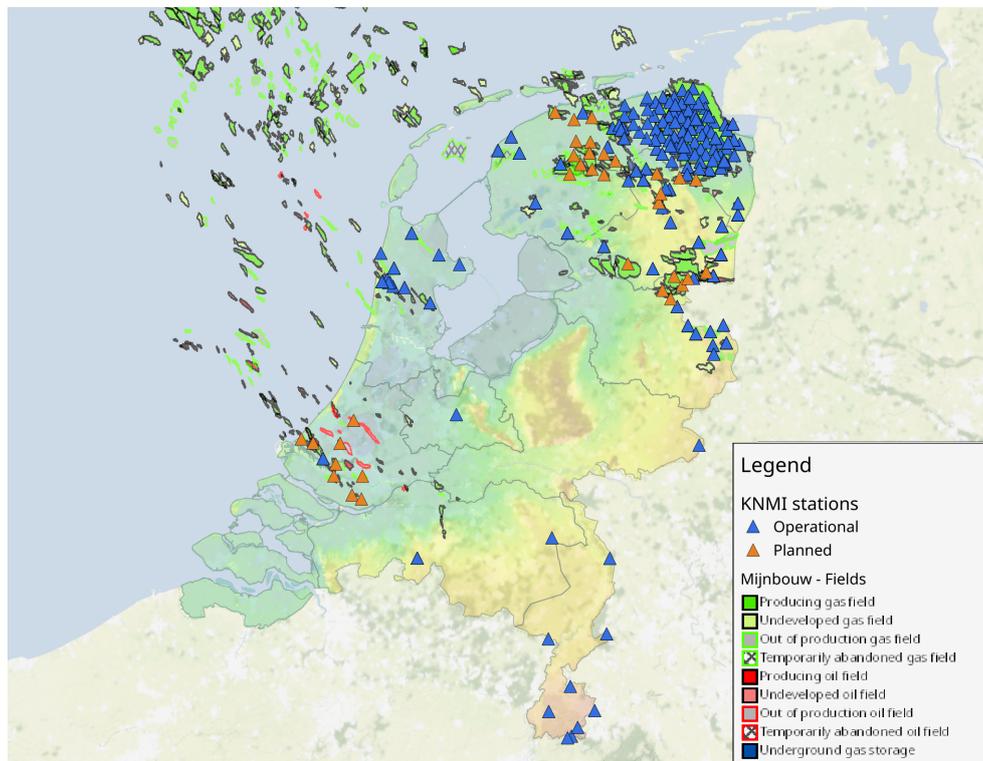
Figure 2.1: Location of active stations (blue triangles), planned stations (orange triangles), and hydrocarbon fields (see legend). At most of the stations, there is one accelerometer or one broadband sensor. There are 123 accelerometers in operation by February 2020 and 16 broadbands. Mining data are from www.nlog.nl.

found in *Ntinalexis et al.* (2019).

Figure 2.1 also shows the KNMI broadband network. Most of these stations are located in the south of the Netherlands where currently no accelerometers are present. Since most of the broadbands have been installed at the Earth's surface, they could also be used to obtain PGV recordings, at least as long as they are not clipped. Besides KNMI, also Utrecht University operates a network of broadband stations in the Netherlands (`http://www.geo.uu.nl/Research/Seismology/nars.html`).

2.2 PGV database

For earthquakes since 2013, the processing of PGV and PGA has been automated at KNMI using the same architecture as developed for the European rapid raw strong-motion database *Cauzzi et al.* (2016). For earthquakes with magnitudes larger or equal to 2.0, the automatically computed values are published on the Rapid Raw Strong Motion portal (RRSM, http://rdsa.knmi.nl/opencms/nl-rrsm/).

The settings for RRSM have been chosen restrictively, to yield robust values without user interaction. As a result, the number of PGV values is relatively low. We compile a new database using less restrictive settings; we include lower magnitudes and records with lower signal-to-noise ratios (SNRs) than within RRSM. In the future, part of the settings will be migrated to RRSM, such that this database can be used for partly automated PGV contour computation.

Since September 2013, EpiSensor accelerometers are in operation that record in a continuous fashion. From the new set of instruments, low-magnitude recordings are available. The data are in general rich in distance and magnitude distribution. However, there are limited data for $M >= 3$ recorded near the epicenters. From 1997 to halfway 2013, GeoSig accelerometers were in the field that recorded ground motion in a triggered mode. Only for large-magnitude events data were retrieved and is still present in KNMI archives. From this archived database, we select all data for $M >= 3$ to supplement the small-distance range. Table 2.1 lists the earthquakes that are used to supplement the PGV database.

Date	Place	Magnitude	Latitude	Longitude
19970219	Roswinkel	3.4	52.832	7.038
19980714	Roswinkel	3.3	52.833	7.053
20001025	Roswinkel	3.2	52.832	7.052
20060808	Westeremden	3.5	53.35	6.697
20081030	Loppersum	3.2	53.337	6.72
20090508	Zeerijp	3.0	53.354	6.762
20110627	Garrelsweer	3.2	53.303	6.787
20120816	Huizinge	3.6	53.345	6.672
20130207	Zandeweer	3.2	53.389	6.667
20130702	Garrelsweer	3.0	53.294	6.785

Table 2.1: Events before September 2013 that are used to supplement the PGV database.

Figure 2.2 shows an example of a 3-component recording $\mathbf{u(t)} = [u_E(t)\ u_N(t)\ u_Z(t)]$ expressed in particle velocity [m/s]. The PGV database is obtained for 3 definitions of PGV. For any definition, the earthquake recording of both horizontal components is taken. Subsequently, for

1. PGVgeo, the geometric mean is taken over the absolute values of the two horizontal components. From this geometric mean trace, the maximum value is stored:
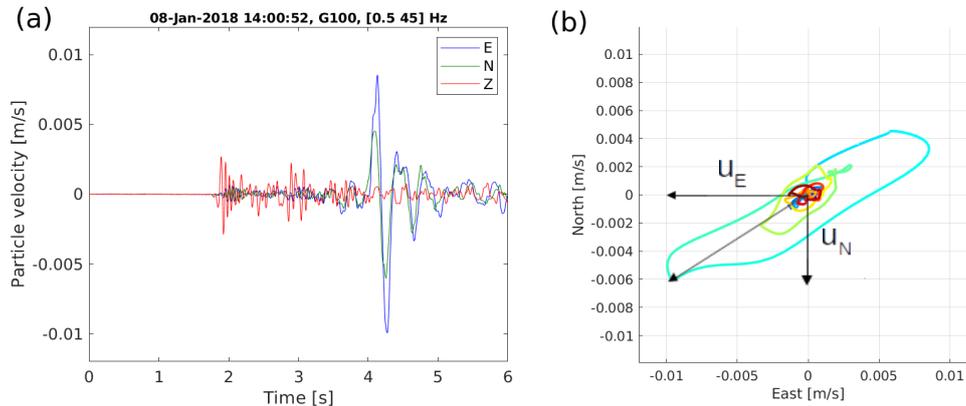
Figure 2.2: (a) An example of a 3-component recording and (b) the corresponding particle motion in the horizontal plane. On (a) the blue, green and red functions are the East, North and vertical-component recordings, respectively. On (b) the colour of the line depicts the timing, going from 3 seconds (dark blue) to 6 seconds (red) with respect to earthquake origin time.

PGVgeo=$max(\sqrt{|u_E(t)||u_N(t)|})$ [1]. For the example in Figure 2.2 this yields 0.0076 m/s.

2. PGVmax, the maximum absolute value is taken over the two components. This value is stored. Hence, PGVmax is a value that was actually measured by one of the two components: PGVmax=$max(|u_E(t)|,|u_N(t)|)$. For the example in Figure 2.2 this maximum is reached at the East component, with a value 0.0100 m/s.

3. PGVrot, the two components are combined to obtain the maximum value in the horizontal plane (by taking the resultant). From this resultant trace the maximum value is taken: PGVrot=$max(\sqrt{u_E^2(t)+u_N^2(t)})$. For the example in Figure 2.2 this yields 0.0114 m/s. The same value is obtained when rotating the recording to any azimuth and storing the results that yield the largest amplitudes, and taking again the maximum amplitude of the resulting trace. For that reason, PGVrot is often referred to as 'maximum rotated'.

The three PGV definitions have the following sequence: PGVgeo≤PGVmax≤PGVrot (e.g., *Bommer et al.*, 2019). Only if both components have the same amplitude, then PGVgeo will be equal to PGVmax and only if one of the components is zero, then PGVmax can be equal to PGVrot. PGVgeo is widely used in earthquake engineering. In the limiting case that the S-waves are perfectly polarized in the direction of one of the components, PGVgeo is equal to zero. PGVmax is an actual ground motion in a direction in which one of the components is oriented. PGVrot is the actual largest horizontal ground motion that materialized.

The PGV database is computed in the three different PGV flavours, to accommodate the demand from the engineering side. Some building codes might be stated in PGVrot, whereas others are defined in terms of PGVmax. Below Table 4.1 conversion factors are given to translate PGVrot to PGVgeo and PGVmax. PGVgeo is used for the comparison with three existing GMPEs (Chapter 3). In all other figures in this report PGVrot is used.

To allow reproducability, below we list the relevant settings for building the PGV database.

---

[1]Sometimes only maximum values per component are available. In this case, PGVgeo is approximated with $\sqrt{max(|u_E(t)|)max(|u_N(t)|)}$, which would yield the same value if the maxima on the two horizontal components occurred at the same time. In our database, PGVgeo is computed directly from the waveforms, using the geometric mean trace.

- We query the KNMI earthquake database (`http://rdsa.knmi.nl/fdsnws/event/1/`) from January 1997 until the end of 2019.

- For each earthquake, all available recordings within a distance range $R_{max} = 6 + 40M$ [km] are downloaded from the KNMI database using FDSN webservices (`http://rdsa.knmi.nl/fdsnws/dataselect/1/`), where M is the local magnitude.

- The upper search yields data from September 2013 from the public database. This is supplemented by data from $M \geq 3$ events from the KNMI archive between February 1997 and July 2013.

- For each station, the relevant information is downloaded from the KNMI inventory (`rdsa.knmi.nl/fdsnws/station/1/`)

- The instrument response is deconvolved with ObsPy (*Beyreuther et al.*, 2010) using the default settings.

- Bandpass filtering.

- Computation of PGVrot (or PGVgeo or PGVmax).

- Retain values with $SNR \geq 6$ dB.

For the band-pass filtering, a fixed band between 0.5 and 45 Hz is taken using Butterworth filters of order 5. Limiting the upper frequency to 45 Hz is done for multiple reasons. The main reason is that risk is not affected by frequencies beyond 45 Hz. Secondly, including frequencies up till 80% of Nyquist (which would be 80 Hz for most sensors) has as a consequence that a few near-field horizontal PGV values are caused by P-waves. We prefer to assess and derive functionals for S-waves only. Thirdly, setting the maximum frequency below 50 Hz has the advantage that no additional notch filter needs to be applied to remove 50 Hz grid distortion (which exist for a part of the stations). A fourth reason is that the older accelerometers had a sampling frequency of only 100 Hz, yielding an upper usable frequency of about 45 Hz.

The SNR is computed on the PGVrot trace. The signal time window is a 5 second time window around the direct S-wave arrival. The noise time window is a 5 second time window from -5 to 0 seconds with respect to earthquake origin time (EOT). The SNR threshold results in rejecting stations where the earthquake ground motion is largely hidden in ambient vibrations. Typically these are stations where the PGV level has dropped below 0.01 mm/s.

For the current database only accelerometers are used. In principle, a part of the broadband stations would be suitable as well since they are also on the Earth's surface and are not clipped for most of the recorded ground motions. However, the sample rate of most broadband stations is only 40 Hz. Therefore they cannot be mixed in with the accelerometer data in a consistent fashion, at least as long the maximum frequency is not further reduced from 45 Hz to about 18 Hz.

No explicit pre-caution is taken for clipping. The accelerometers in the network have different clipping levels: $4g$ (for the B-network) and $2g$ (for all other networks). At 5 Hz, $4g$ and $2g$ corresponds to 125.0 and 62.5 cm/s, respectively. This is significantly higher than the largest PGVrot value in the database, which is 5.61 cm/s. This value was recorded at station ROS1 during the Roswinkel 19-02-1997 event.

First we query the database for events with $M \geq 1.2$. 233 onshore events are found from September 2013. From these events, there are 200 events that have at least one accelerometer recording with sufficient SNR. These events are supplemented with data from the 10 older events as listed in Table 2.1. For these 210 events PGVrot is computed and for each
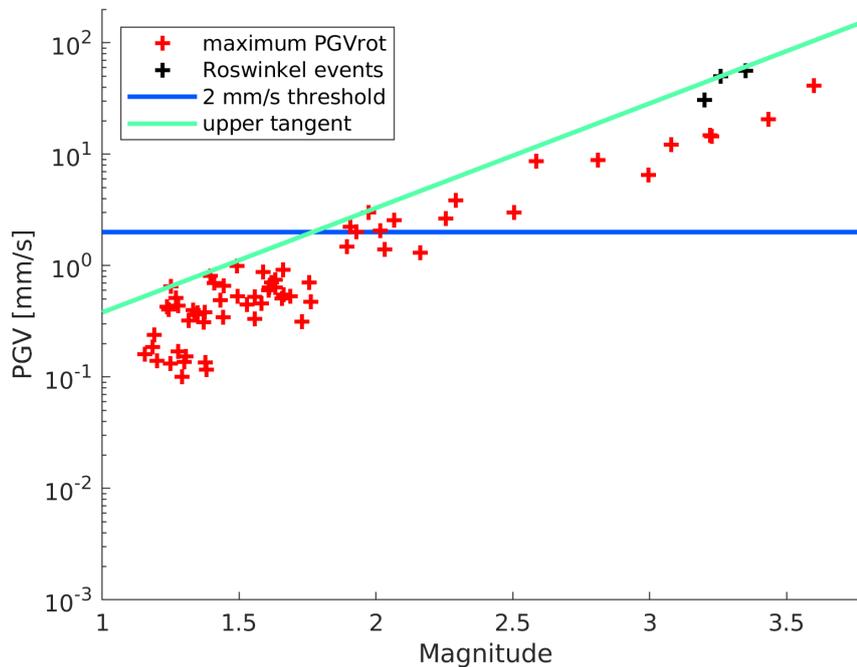
Figure 2.3: Maximum PGVrot values per event, for 71 events from the onshore database of induced events between 1997 and 2019, with a mimimum magnitude of 1.2 and with a recording within 2 km epicentral distance. The horizontal colored line denotes the lowest PGV threshold level. The upper tangent line to the data points yields an estimate of the dPGV/dM derivative.

event only the largest PGVrot is retained. If this PGVrot was measured at an epicentral distance within 2 km, it is plotted as function of magnitude. Figure 2.3 shows the resulting 71 data points. Using this figure, an empirical estimate is made of the minimum relevant magnitude. For a PGVrot threshold of 2 mm/s, the first recording of this level has occurred with a $M{=}1.9$ event.

The maximum PGV records per event (Figure 2.3) is also used to find an estimate of the magnitude scaling of ln(PGV). The upper tangent line is an estimate of the maximum possible PGV levels as function of magnitude. The slope of this line is 2.16 expressed in ln(PGV) per magnitude.

In the following we will restrict the database to a minimum magnitude of 1.5. Figure 2.4 shows all the onshore induced events for $M \geq 1.5$ in the database for which at least one good-quality (SNR$\geq$6 dB) accelerometer record is available. In total there are 114 events, of which 11 are outside the Groningen field.

2.3 Radiation pattern

An additional source of ground-motion information is from moment tensors. Isotropic sources radiate energy equally to all directions and hence lead to quite symmetric and smoothly varying ground motions. Actual earthquakes can be largely described as pure slip on a fault. Such a double-couple source has an anisotropic radiation pattern. This leads to
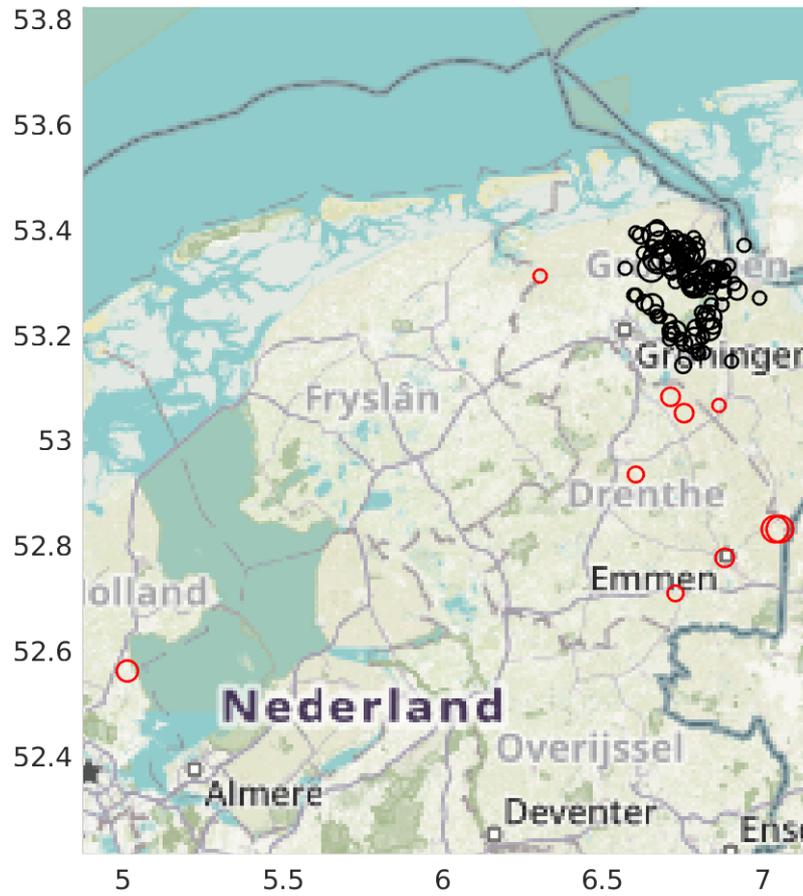
Figure 2.4: Map showing events within the PGV database for $M \geq 1.5$. 103 events (black circles) are related to the Groningen gas field, the remaining 11 events (red circles) are related to other gas fields.
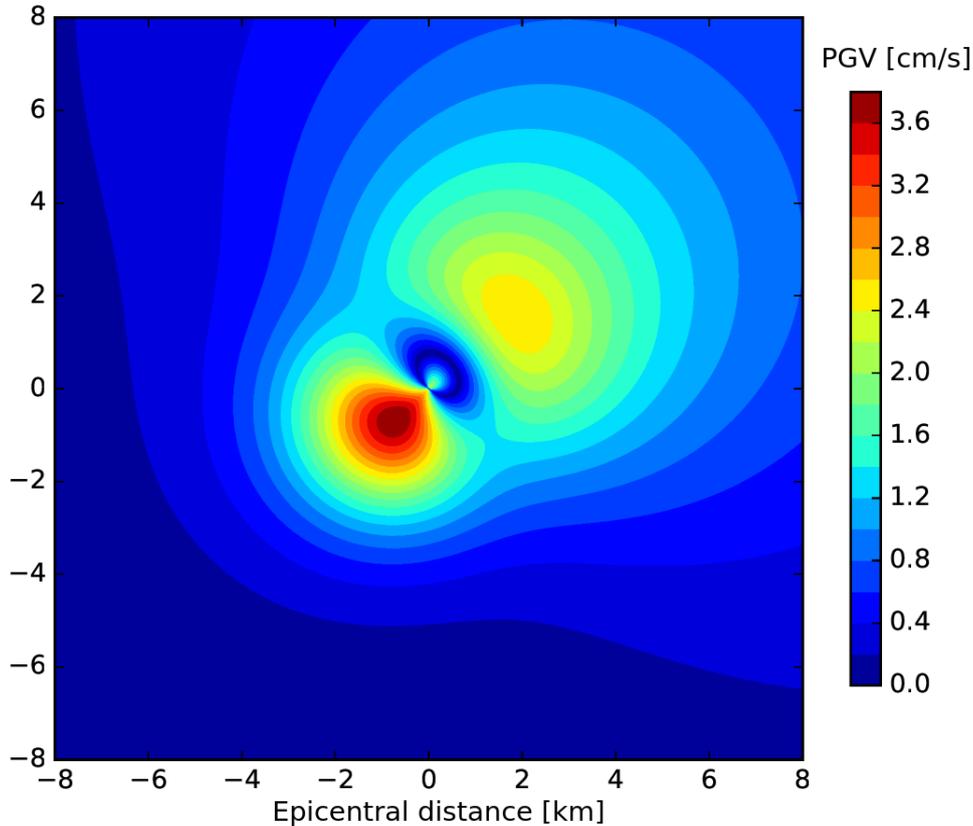
Figure 2.5: PGV in cm/s at the surface due to a source at 3 km depth in Groningen. The source mechanism is strike 315, dip 62 and rake -95 degrees (Zeerijp event 08-01-2018, M 3.4). The motion modeled is the total shear wave motion, vector summation of the SV and SH movement. For conversion from displacement amplitudes to velocity, a dominant frequency of 3 Hz has been assumed. Geometrical spreading was implemented as $R^{-1.3}$, comparable to average attenuation used in local magnitude calculations for Groningen. An averaged Groningen velocity model was used in the calculations.

quite strong variations in ground motion.

Induced seismic events in the Netherlands are characterized by normal faulting (*Willacy et al.*, 2019) and a negative isotropic component (*Dost et al.*, 2020). Dips for events analyzed in Groningen vary between 50 and 70 deg. Similar normal faulting is expected for events outside of Groningen.

Figure 2.5 shows the radiation pattern for the Zeerijp 08-01-2018 event, projected to the Earth's surface. The algorithm used is described by *Ebel and Bonjer* (1990). A strong variation of ground motion can be observed. The peak in ground motion has an offset of about 1 km with respect to the epicenter (which is at 0.0 km). Moreover, 3 lobes can be distinguished: strong motions in the SW of the epicenter, a lobe with much smaller amplitudes directly NE of the epicenter and again a lobe with stronger amplitudes further to the NE.

On Figure 2.6 the resulting PGV as function of epicentral distance is shown. Here again it can be identified that the peak PGV occurs at about 1 km. Moroever, it can be seen that near the source there is a very strong variation of amplitudes. In reality, this variation will be reduced by the arrival of (strong) secundary waves which have not been included in the
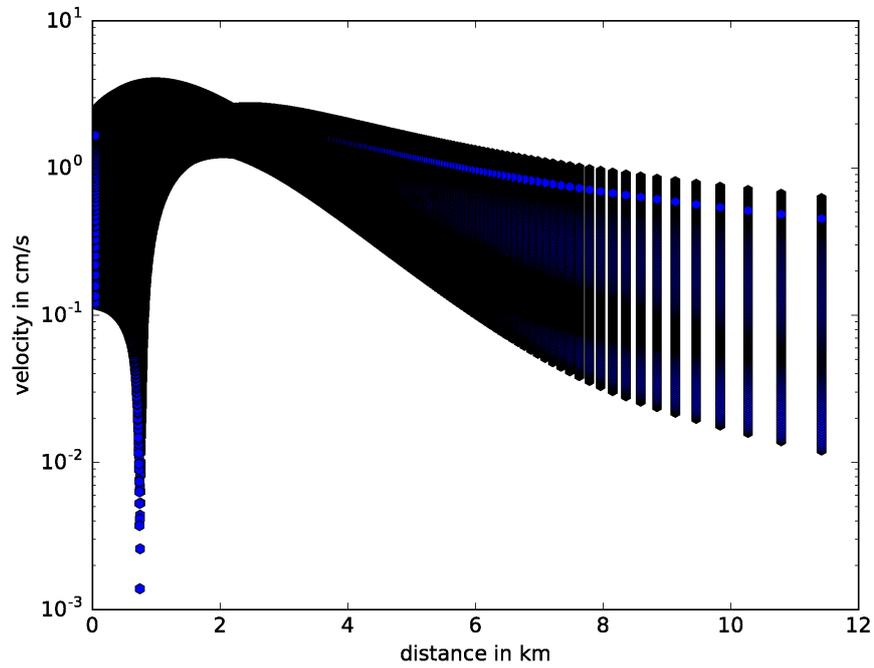
Figure 2.6: PGV variations with distance from the epicenter. All grid points used in Figure 2.5 are shown. The values are shown as blue circles with black bounding lines. Where the density of circles is high, they appear fully black.

modeling.

3 Existing GMPEs

For computing PGV threshold regions in areas where no local recordings exist we need to resort to ground-motion prediction equations (GMPEs). GMPEs have been derived for many different tectonic- and induced-seismicity settings. Instead of deriving a new GMPE specifically for induced events outside Groningen, we first assess whether existing GMPEs already well capture the mean observed PGV levels. Moreover, the distribution of the PGV recordings needs to be well described by the listed standard deviations. In the following we list the functional forms and variability for 3 GMPEs that have been constrained wholly or partly with Dutch data and that were derived for seismicity in the upper crust. Additionally, we consider another recent GMPE that has been constrained with many shallow tectonic events. In the next chapter, these GMPEs are tested against the PGV database.

The aleatory variability in the models is expressed as a standard deviation. This is the standard deviation of the residuals between the measurements and model predictions. The model itself is calibrated to predict the median motion (using the logarithmic mean of the recordings). Only if the residuals are normally distributed, and a large number of local recordings is used for calibrating the model, the standard deviation would be an accurate description of the uncertainty. However, when e.g. a certain model is used for another region, with a different faulting style, its printed standard deviation might not be sufficiently descriptive for this other region. No alternative models are described to account for epistemic uncertainty.

3.1 Dost

The Dost GMPE has been published in *Dost et al.* (2004). It is based on an attenuation relation derived for the north of the Netherlands, which has been derived for the purpose of magnitude computation with geophones installed at depth. This attenuation relation has been scaled with recordings at the Earth's surface at three source regions in or near the Netherlands: Voerendaal, Alsdorf and Roswinkel. All three regions have seismicity at depth without an overlaying salt layer. These source regions have different depths and hence a depth dependency has been introduced. The relation reads

$$\log(Y_{mod}) = -0.53 + 0.74M - 0.00139r - 1.33\log(r), \qquad (3.1)$$

where r is the hypocentral (source-to-site) distance in km, M is the local magnitude and Y_{mod} is the geometric mean over both horizontal components in mm/s[1].

In this model, no saturation has been built in for higher magnitudes. The model was obtained from earthquakes with magnitudes between 0.8 and 4.9. Most of the used strong motion recordings are from within 5 km distance. However, the underlying attenuation re-

---

[1]In *Dost et al.* (2004) Y_{mod} is expressed in cm/s and hence the first term on the right-hand side of the original equation reads -1.53.

lation is from a much wider epicentral distance range. The total model standard deviation σ_{mod} reads 0.33, expressed in *log*, or 0.76 expressed in *ln*.

3.2 Douglas

Douglas et al. (2013) derived a GMPE specifically for geothermal areas. Their model is based on data from 6 source areas. Besides from 4 geothermal areas, data are included from natural seismicity (Voerendaal) and from gas extraction (Roswinkel). They used data with focal depths down to about 10 km and hypocentral distances until 50 km. They derived a model including and excluding site effect. Their mean PGV model including (average) site effects reads

$$ln(Y_{mod}) = -3.459 + 2.018M_w - 1.124ln(\sqrt{r^2 + 2.129^2}) - 0.046r, \qquad (3.2)$$

where ln is the natural logarithm, M_w is the moment magnitude, r the hypocentral distance in km and Y_{mod} is the peak ground velocity obtained by taking the geometric mean over both horizontal components in mm/s^2.

The variability in the model is large, probably due to combining data from different type of sources and largely varying subsurface and site conditions. The within-event variability ϕ_{mod}, the between-event variability τ_{mod} and total variability σ_{mod} are 1.811, 0.745 and 1.958, respectively. These are variabilities expressed in natural log ln.

3.3 Bommer

The Bommer model (*Bommer et al.*, 2019) is parameterized as

$$ln(Y_{mod}) = c_1 + c_2M + g(R^*), \qquad (3.3)$$

where M is the local magnitude, $g(R^*)$ is a function that describes the geometrical spreading and the effective point-source distance R^* is defined as follows

$$R^* = \sqrt{R^2 + [e^{0.4233M - 0.6083}]^2}, \qquad (3.4)$$

where R is the epicentral distance in km. Equation 3.4 expresses the magnitude-dependent distance saturation term.

The geometrical spreading is divided in three distance ranges with different functional forms:

$$g(R^*) = c_4 \ln(R^*) \qquad R^* \leq d_1$$
$$g(R^*) = c_4 \ln(d_1) + c_{4a} \ln(R^*/d_1) \qquad d_1 < R^* \leq d_2$$
$$g(R^*) = c_4 \ln(d_1) + c_{4a} \ln(d_2/d_1) + c_{4b} \ln(R^*/d_2) \qquad R^* > d_2. \qquad (3.5)$$

The GMPE has been derived for three different definitions of *PGV*, leading to different numerical values for the coefficients. The distances d_1 and d_2 do not vary for different definitions of PGV. All parameters are listed in Table 3.1.

---

[2]In *Douglas et al.* (2013) Y_{mod} is expressed in m/s and hence the first term on the right-hand side of the original equation reads -10.367.

Coefficient	PGVgeo	PGVmax	PGVrot
c_1	-3.2907	-2.8979	-2.7738
c_2	2.24816	2.28589	2.2835
c_4	-1.75493	-1.90988	-1.93283
c_{4a}	-1.14046	-1.11959	-1.10756
c_{4b}	-1.61257	-1.65679	-1.67393
d_1 [km]	6.32	6.32	6.32
d_2 [km]	11.62	11.62	11.62

Table 3.1: Parameters of the Bommer empirical model (*Bommer et al.*, 2019)

With respect to the original c_1 coefficients in *Bommer et al.* (2019) we added the term $ln(10) = 2.3026$ to obtain the values as listed Table 3.1. This term serves to convert from cm/s, which is used in the reference, to mm/s, which is used in this report.

The Bommer model has an empirical base between M=1.8 and M=3.6 and for distances up till 35 km. Note that the GMPE has been derived with Groningen data only. Hence, for a reservoir below a thick layer of salt at around 3 km depth. Variabilities in the model are split up into within-event variability ϕ_{mod} and between-event variability τ_{mod}. This yields the total standard deviation of the model $\sigma_{mod} = \sqrt{\phi_{mod}^2 + \tau_{mod}^2}$. The variabilities are expressed in ln and are, for the different definitions of PGV listed in Table 3.2.

Coefficient	PGVgeo	PGVmax	PGVrot
τ	0.25128	0.25169	0.25242
ϕ	0.48205	0.54001	0.53613
σ	0.54361	0.59578	0.59258

Table 3.2: Total σ, intra-event ϕ and inter-event τ standard deviations in the Bommer empirical model.

A widely applied proxy for site amplification is the average shear-wave velocity over the top 30 meters: V_{S30}. In a previous version of the model (*Bommer et al.*, 2016) V_{S30} was explored as an additional variable within the GMPE. Since the influence was found to be small, it was omitted to yield a simplified model.

The Bommer empirical model that is used in this report is not to be confused with the Bommer ground motion model that has been developed for hazard and risk computations (*Bommer et al.*, 2017). The latter model includes site terms and extrapolation to events with magnitudes much higher than observed. Both elements are not present in the Bommer empirical model.

3.4 Atkinson

Atkinson (2015) published a preliminary GMPE for small to moderate size events (M=[3 6]) at short hypocentral distances (< 40 km) and refined the model later for events of M=[3 5] at distances within 5 km of the epicenter (*Atkinson et al.*, 2016). Especially the last publication is of interest to this study because it uses a large database of induced events recorded at close distances in the Geysers region (California, USA). In *Atkinson* (2015), model results have been compared with data from *Douglas et al.* (2013), part of which consists of data from Roswinkel.

The model is parameterized as

$$ln(Y_{mod}) = -3.151 + 1.762M - 0.09509M^2 - 1.669logR^*, \qquad (3.6)$$

Similar to the Bommer model, R^* in the above equation is the effective point source distance, which is defined as

$$R^* = \sqrt{r^2 + [max(1, 10^{(-0.28+0.19M)})]^2}. \tag{3.7}$$

A difference with Bommer is that the hypocentral distance r [km] instead of the epicentral distance R appears in the effective distance term.

With respect to the original coefficients in *Atkinson* (2015) we added the term $log(10) = 1$ to obtain equation 3.6. This term serves to convert from cm/s, which is used in the reference, to mm/s, which is used in this report.

The model is based on the geometric mean of the horizontal components. The within-event variability ϕ_{mod}, the between-event variability τ_{mod} and total variability σ_{mod} are 0.28, 0.18 and 0.33, respectively. These are variabilities expressed in *log*. Expressed in *ln* these values would be 0.64, 0.41 and 0.76.

At epicentral distances > 10 km, the Bommer and Atkinson models are very similar in form and can be matched by adapting the coefficients in equation 3.6. At shorter distances the Atkinson model shows a smoother increase in amplitude towards lower distances.

4 GMPE selection

In this chapter we evaluate which GMPE (Chapter 3) fits best with the PGV database (Chapter 2.2). For the comparison we distinguish 2 different subdatasets: 1) events related to the Groningen gas field and 2) events outside the Groningen gas field. In the second dataset, the PGV distribution is limited in magnitudes and distances. For that reason, before comparison with GMPEs, all recorded PGV values are first remapped to a magnitude 2.2 event.

In the previous chapter, we derived from the data a magnitude scaling of 2.16 ln(PGV) per magnitude. This value is considerably larger than in Dost (1.71) and larger than the value listed in Douglas (2.018). In Bommer, the derivative is distance dependent. It varies from 1.51 at R=0 to 2.25 at large distances. In Atkinson, this derivative is magnitude and distance dependent. For R=0 and the magnitude range of interest(M=[1.5 3.6]) the derivative goes down from 3.3 to 2.3. We use our estimate of the magnitude scaling to map all data points to a magnitude 2.2 earthquake. This estimated magnitude scaling is merely used to facilitate an easy comparison of the database with the different models. We do not recommend this scaling for usage in a GMPE.

In the last section, candidate models are adapted to further improve the misfit with the PGV database.

4.1 GMPE for events in Groningen

Figure 4.1 shows distributions for the within-Groningen PGV dataset. In total there are 3192 PGV values. Only a small proportion is from the period 2006 until halfway 2013, when only the triggered network was operational. Most of the recordings are from between 2015 and 2019. In October 2014, the first batch of stations of the G-network (*Dost et al.*, 2017) were delivering data. By the end of 2017 this new network was fully operational.

Most of the recordings are within 40 km from the epicenter. However, also some recordings between 80 and 140 km distance passed the SNR threshold. Figure 4.2 shows all PGV values mapped to a magnitude of 2.2, in comparison with the 4 GMPEs. It can be seen that most values straddle around the Bommer model. The values beyond 80 km distance are somewhat overestimated with this GMPE. This distance range is far beyond the 40 km range for which this GMPE was calibrated.

Figure 4.3 shows the PGV recordings up till a distance of 20 km. The recorded PGV values are almost all below the Dost model, and almost all above the Atkinson model. The Douglas model is quite a good mean between about 5 and 20 km. At shorter range, most PGV values are above the Douglas model. The Bommer model, on average, describes the PGV values best.
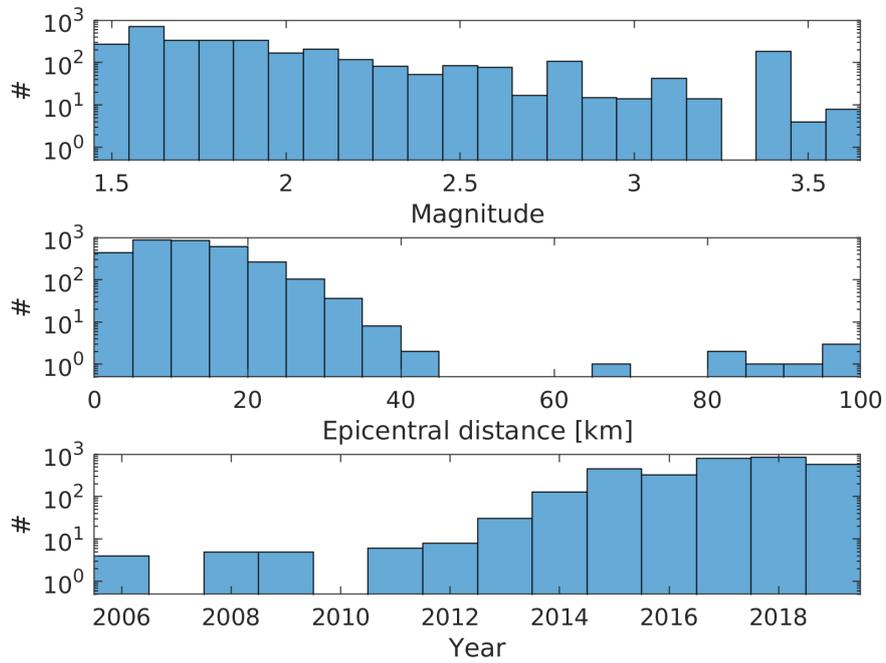
Figure 4.1: PGV database distribution in respect of magnitude, distance and year, for events within Groningen. The numbers are shown on a log scale.
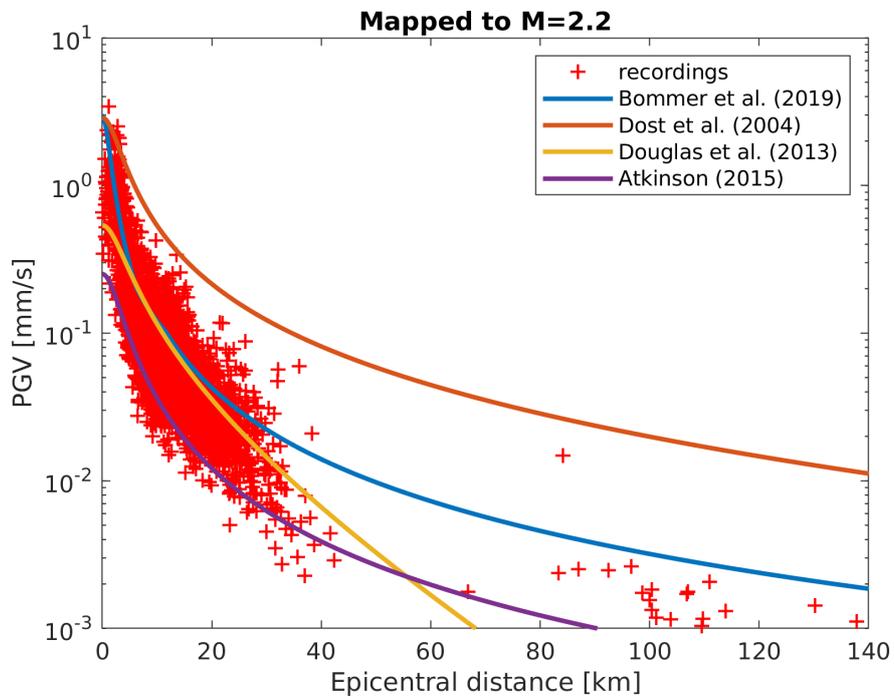
Figure 4.2: PGVgeo values for events within Groningen and comparison with 4 GMPEs.
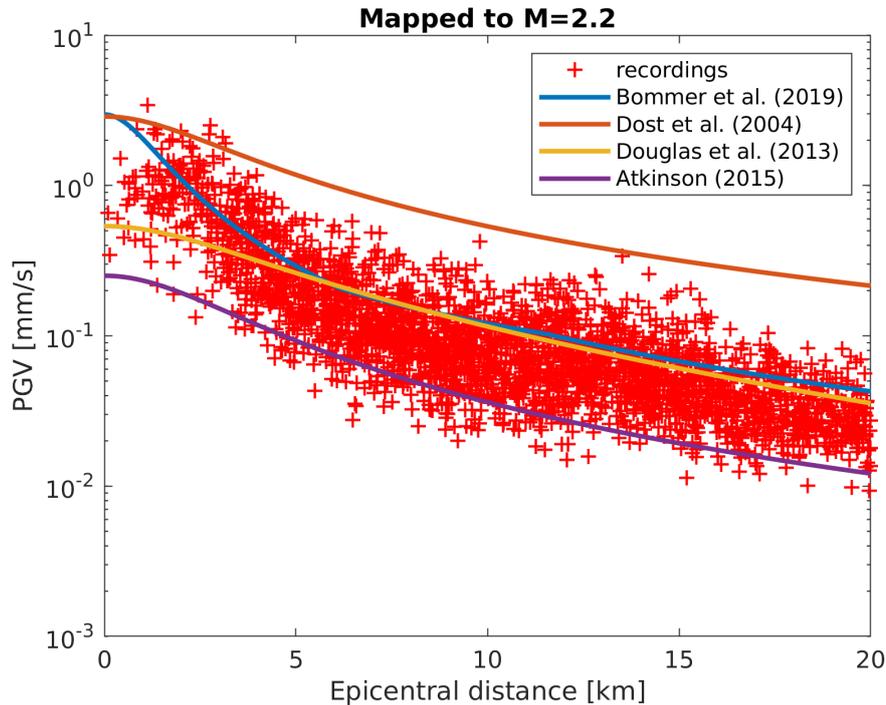
Figure 4.3: PGVgeo values for events within Groningen and comparison with 4 GMPEs, for distances up till 20 km.

In the same figure it can be noted that Dost, Bommer and Atkinson have very similar attenuation characteristics at larger distances. The distance attenuation at short range, however, is very different in these three models.

4.2 GMPE for events outside Groningen

Figure 4.4 shows distributions for the outside-Groningen PGV dataset. In total there are 99 PGV values. Most of them are from 2016. All data from before 2013 are from Roswinkel. These Roswinkel events are also responsible for all data points at magnitudes larger than 2.5.

Figure 4.5 shows the PGV recordings up till a distance of 90 km. The recorded PGV values are all below the Dost model and almost all above the Atkinson model. The Douglas model is quite a good mean between about 5 and 30 km. At shorter and further range, the Bommer model better describes the attenuation. Also at intermediate range, the Bommer model does a reasonable job.

Figure 4.6 shows the PGV values together with the mean Bommer model and different confidence regions. It can be seen that the data are on average lower than predicted by the Bommer model. This difference is taken care of when computing the event term (Section 5.1). The variability in the Bommer model, which dictates the size of the confidence region, matches well with the data distribution. If the model were shifted downwards, all points would map within the 95% confidence zone.

Figures 4.5 and 4.6 show that also outside Groningen, the Bommer GMPE best describes
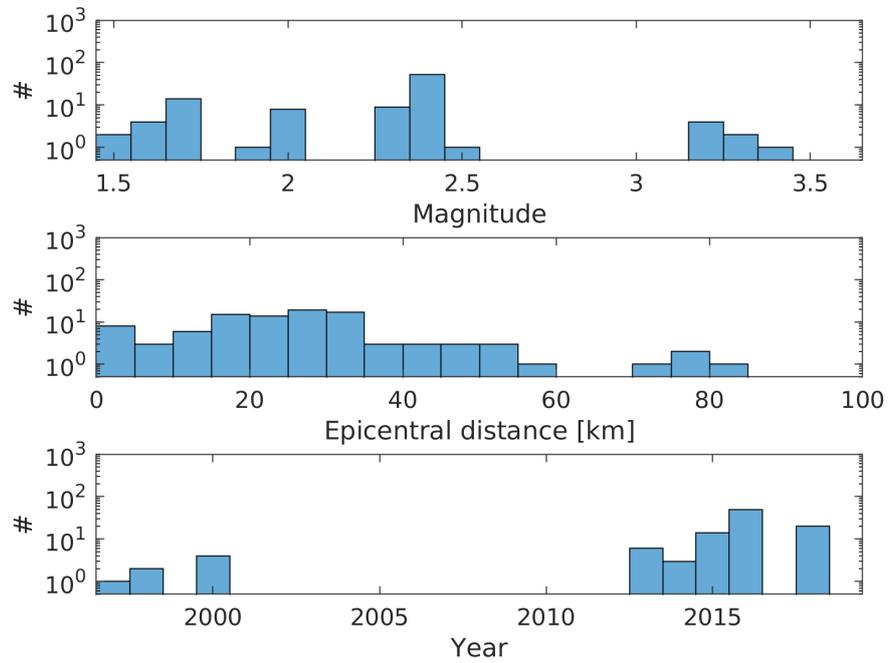
Figure 4.4: PGV database distribution in respect of magnitude, distance and year, for events outside Groningen. The numbers are shown on a log scale.
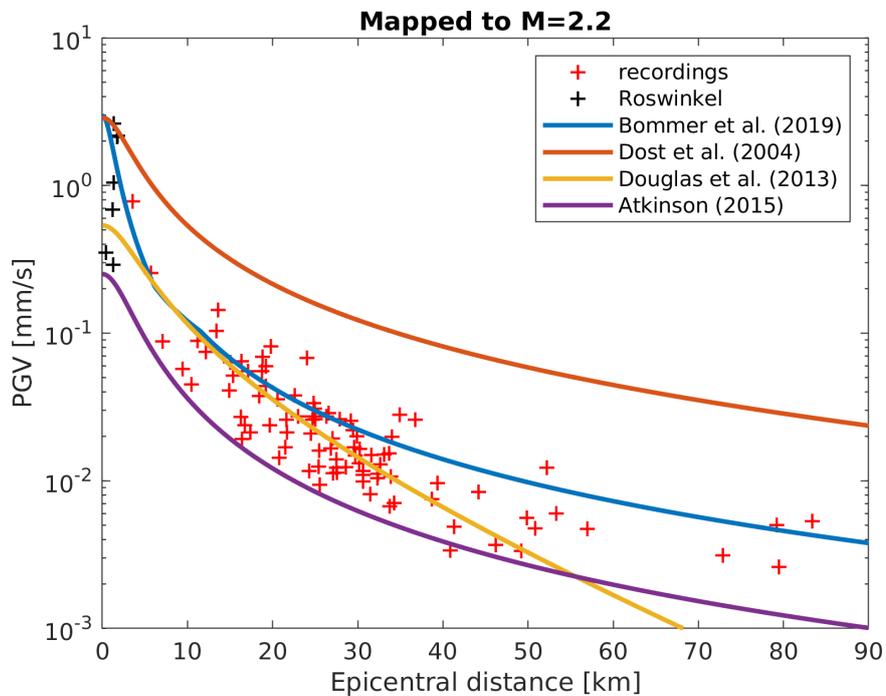
Figure 4.5: PGVgeo values for events outside Groningen and comparison with 4 GMPEs.
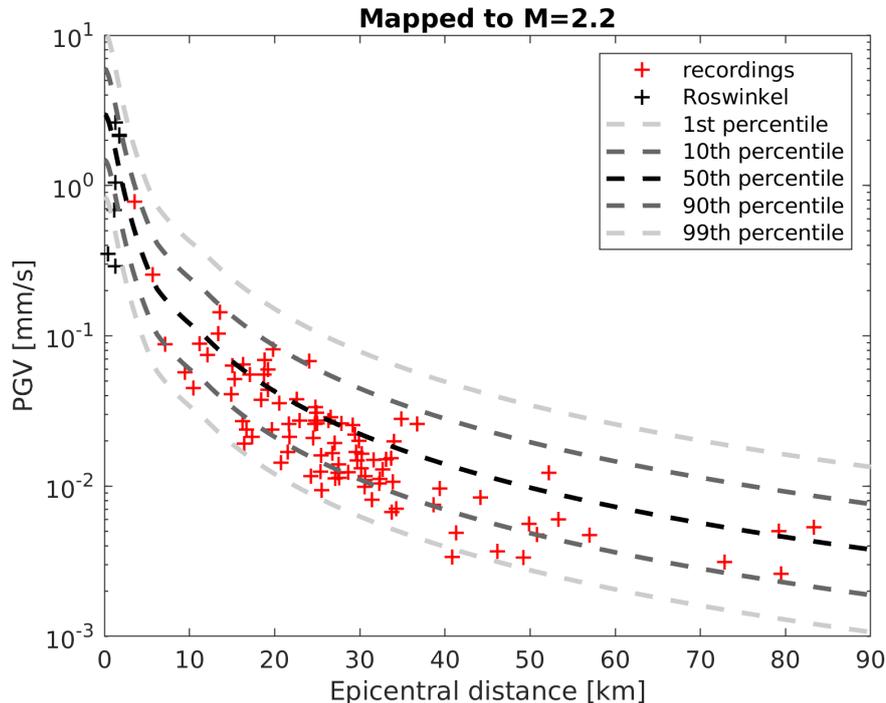
Figure 4.6: PGVgeo values for events outside Groningen and comparison with the mean Bommer model (50th percentile) and uncertainty bounds (all other percentiles).

the attenuation for induced earthquakes. This might not be surprising, since most gas fields have a similar setting as the Groningen gas field: a reservoir sealed by a thick layer of salt. And also the near-surface outside of Groningen is characterized by a similar mix of unconsolidated sediments. An exception to this are the Roswinkel events, as depicted with black crosses in the figure. The Roswinkel reservoir is not sealed by Zechstein evaporites and is shallower (2 km deep) than the Groningen reservoir (3 km deep). Yet, the Roswinkel PGVs do not show a markedly different behavior compared to the other PGV values. A likely reason for this is that the different setting of the hypocenters is already reflected in the magnitude computation. The Roswinkel event with most recordings is analyzed in more detail in Section 6.3.

Note that the above analysis was done for a limited number of 99 PGV values which remain to have a strong bias towards seismicity and site conditions in the northeast of the Netherlands. With the future expansion of the 'outside Groningen' database, the above analysis can be further refined.

4.3 Misfit for Bommer model

In the previous sections we found that induced events within and outside Groningen have similar attenuation characteristics. Therefore, the most recent GMPE that was derived with a rich dataset of Groningen PGVs (*Bommer et al.*, 2019) was found to be the most suitable model both in and outside Groningen. In this section, we compute the distance-dependent misfit between the complete PGV database (Section 2.2) and the Bommer model.

Figure 4.7 shows the misfit between the PGV database and the Bommer GMPE. For all magnitude ranges, except $3.0 < M \leq 3.6$, the GMPE overestimates the recordings over the
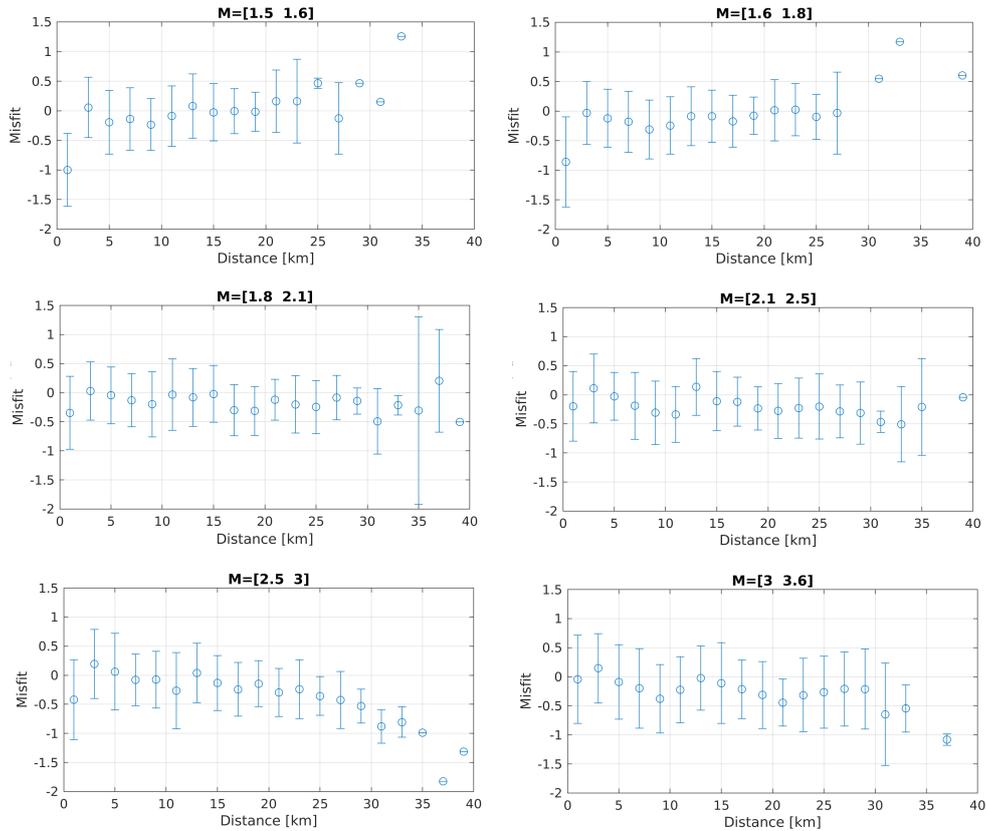
Figure 4.7: Misfit between PGV database (including events within and outside Groningen) and Bommer GMPE for different magnitude ranges. The mean misfit (circles) and standard deviation (the bars denote from -1 to +1 standard deviation) are displayed over distance bins of 2 km.

first few kilometers, leading to a negative misfit (=observation minus model). This effect is most pronounced for magnitudes $1.5 < M \leq 1.8$, i.e., outside the range for which the model was developed. Also between about 5 and 10 kilometers distance, the model overestimates the observed ground motions. At larger distances, the mean misfit is close to zero for all magnitude ranges, and the standard deviation of the misfit is quite constant. For distances larger than 30 km, the misfit is irregular. This is related to the limited number of observations at these larger distances.

Especially the large misfit of the Bommer model at short range, motivates us to look at alternative models or adaptations of Bommer, as introduced and tested in the next section.

4.4 Testing adapted models

In this section, we propose modifications to the Bommer and Atkinson models and test whether these modifications lead to a better description of our dataset. Since the data from events within and outside Groningen are not markedly different (Figures 4.2 and 4.5) we do the testing on the complete PGV dataset. In the text and figures, results for PGVrot are shown. However, the models can be converted to PGVgeo and PGVmax.

The adapted models need to perform well especially in the distance range in which 2 mm/s

is crossed. In the Bommer model, for $M=3.6$, for P50, P90 and P99, these distances are 15.7, 24.8 and 36 km, respectively. Beyond 30 km, the number of recordings goes down (Figure 4.1) and the computation of fitting statistics is compromised (Figure 4.7). Taken these two aspects in consideration, we test the different model adaptations up till distances of 30 km.

The adapted model that is being tested as the best one, will be used in the remainder of the report for the computation of PGV threshold regions.

4.4.1 Bommer adaptation

A drawback of the Bommer model is that it has been derived for Groningen seismicity only, with a event depth around 3 km, which makes less applicable for regions with largely different focal depths. To also directly mend this issue we introduce the depth D into the effective distance term, yielding

$$R^* = \sqrt{R^2 + D^2 + [e^{e_1 M + e_2}]^2},$$ (4.1)

in which the first two terms under the square-root are the hypocentral distance squared: $r^2 = R^2 + D^2$.

The addition of D^2 makes the distance-saturation term considerably larger. This increase cannot be completely undone by fitting new values for e_1 and e_2, since the exponential function will remain positive (as long as it is real valued). Thus, after the introduction of a hypocentral distance, also other coefficients in the model need to be re-calibrated. We retain the magnitude-scaling coefficient c_2 and re-estimate all other coefficients using the PGV database. In the fitting, we use as a constraint to retain the attenuation at large distances. This attenuation had been derived in *Dost et al.* (2004) and worked well until 40 km distance in *Bommer et al.* (2019). Table 4.1 contains the parameters for the adapted model, which is coined BMR2. The parameters e_1 and e_2 appear in equation 4.1. The coefficients c_1, c_2, c_4, c_{4a} and c_{4b} and the distances d_1 and d_2 can be found in equations 3.3 and 3.5. For the BMR2 model we retain the variability as computed for the Bommer empirical model (Table 3.2).

Coefficient	PGVrot
c_1	2.2800
c_2	2.2835
c_4	-4.2800
c_{4a}	-0.8000
c_{4b}	-1.7000
e_1	0.0600
e_2	1.1300
d_1 [km]	8.10
d_2 [km]	11.62

Table 4.1: Parameters of the adapted Bommer model, which is coined BMR2.

The above model is based on the PGV database for $M \geq 1.5$ (Section 2.2). For that same database, conversion factors can be derived between different definitions of PGV. This allows mapping the adapted models to PGVgeo and PGVmax. For going from PGVrot to PGVgeo the mean and the standard deviation are 0.6074 ± 0.0934. Note that this distribution has a large negative skew: -1.1270. Conversion from PGVrot to PGVmax could be done with 0.9218 ± 0.0733. This distribution also has a negative skew: -0.9008. For deriving conversion factors, we assumed that these are not distance dependent.

Figure 4.8 shows the PGV versus magnitude behavior for the Bommer and BMR2 models, for a source at 3 km depth, predicted at 0 km epicentral distance. At this distance, PGV levels in the BMR2 model are lower for most magnitudes. Only at $M \geq 3.6$ the BMR2 model has larger values.

The parameter e_1 in Table 4.1 is fairly small. As a result, there is a weak magnitude dependence in the effective distance term (equation 4.1). This yields a fair fit to the PGV database, for the considered magnitude range. However, as is shown e.g. with stochastic modelling (*Douglas and Jousset*, 2011) it is likely that for larger magnitudes a stronger saturation is needed, in order to reduce the magnitude scaling.

4.4.2   Atkinson adaptation

Since the Bommer model fits the data at longer distances well, we will first compare the Atkinson model to the Bommer model. We have to take into account that the Atkinson model is only valid for the PGVgeo. At epicentral distances > 10 km, the Bommer and Atkinson models are very similar in form, but with an offset between the curves. At shorter distances the Atkinson model shows a smoother increase in amplitude towards lower distances. Therefore, we calculated the differences between the models as a function of magnitude and fit the difference according to the parameterisation used by Atkinson. In addition to the PGVgeo, we also calculated the difference with respect to PGVrot in order to explore if the latter requires a significantly different relation. We determined for a number of magnitudes the best estimate of parameter for the first 3 coefficients in the Atkinson GMPE. For PGVrot, this yields

$$ln(Y_{mod}) = -1.2660 + 0.9973M - 0.0010M^2 - 1.669logR^*, \quad (4.2)$$

and for PGVgeo

$$ln(Y_{mod}) = -1.3327 + 0.9739M - 0.00037M^2 - 1.669logR^*. \quad (4.3)$$

The modified Atkinson model, which is coined ATK2, provides a smoother behavior at small distances. However, in all magnitude bins there is a slight underprediction at these small distances. This can be tuned by adding an additional geometric spreading term, as was discussed in *Atkinson et al.* (2016). In order to lower the misfit in the first 5 km, a geometrical spreading factor $g(R^*)$ is added to the Atkinson model, which factor is optimized for different magnitude ranges:

$$g(R^*) = 0.005 \exp(R^{*-1.3}) \quad M < 1.8$$
$$g(R^*) = 0.014 \exp(R^{*-1.3}) \quad 1.8 \leq M \leq 2.1$$
$$g(R^*) = 0.033 \exp(R^{*-1.3}) \quad M > 2.1. \quad (4.4)$$

Figure 4.9 shows the Bommer (solid lines), BMR2 dotted lines) and ATK2 GMPE (dashed lines) for the magnitude range in which the model is used in this report. Bommer derived their model for magnitudes between 1.8 and 3.6. The adapted models show the largest changes in the first 10 km. ATK2 has much lower values than Bommer in the first 4 km and then has higher values until about 10 km epicentral distance (for a source at 3 km depth). Over the first 5 km, BMR2 has higher values than ATK2, but mostly lower values than Bommer. An exception to this is at high magnitudes, where BMR2 has the highest amplitudes at short range. Between 6 and 10 km, BMR2 has lower amplitudes than Bommer and ATK2.
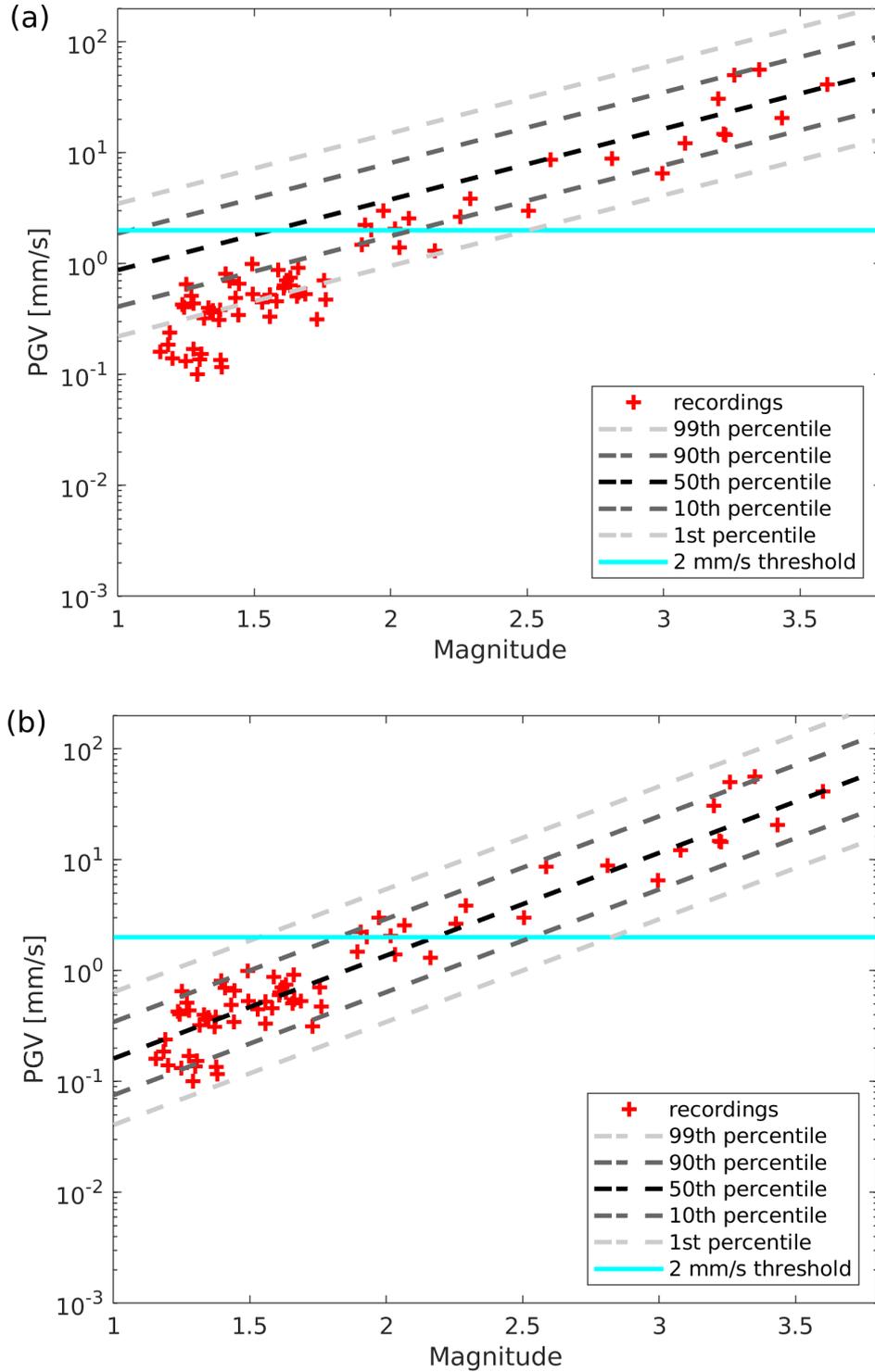
Figure 4.8: Maximum PGVrot values per event (red stars) together with different percentiles (dashed lines) of the (a) Bommer empirical and (b) the BMR2 model. The model values are shown for a distance of 0 km. For computing the percentiles, the total variability σ_{mod} is used. The PGV maxima are shown for 71 events from the onshore database of induced events between 1997 and 2019, with a minimum magnitude of 1.2 and with a recording within 2 km epicentral distance. The horizontal colored line shows one PGV threshold level.
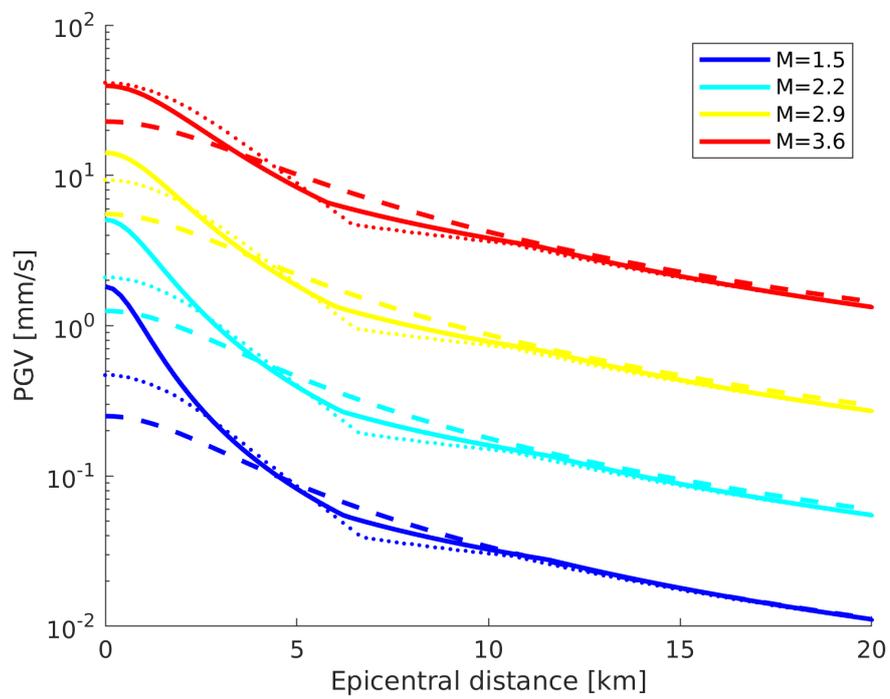
Figure 4.9: three different GMPE attenuation curves for 4 different magnitude classes for epicentral distance up till 20 km; the Bommer model (solid lines), the Bommer model after rewriting the model to hypocentral distance and re-calibrating coefficients (BMR2 model, dotted lines) and the adapted Atkinson model (ATK2 model, dashed lines).

4.4.3   Model testing

For the different models we determine the root-mean-square error:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\{ln(Y_{obs,i}) - ln(Y_{mod,i})\}^2},\qquad(4.5)$$

where $Y_{obs,i}$ is the observed PGV, $Y_{mod,i}$ the modeled PGV and n is number of recordings. For each model, we determine the RMSE for the entire dataset (for $R \leq 30$ km) and for the different magnitude classes. The results are shown in Table 4.2. The RMSE is computed without implementation of event terms. Since the models are meant to represent the mean of the datapoints, the RMSE is practically identical to the model standard deviation.

Magnitudes	[1.5 1.6]	[1.6 1.8]	[1.8 2.1]	[2.1 2.5]	[2.5 3.0]	[3.0 3.6]	[1.5 3.6]
Data points	510	818	914	288	290	242	3062
Bommer	0.5118	0.5182	0.5156	0.5133	0.5155	0.5225	0.5340
BMR2	0.5087	0.5084	0.5130	0.5121	0.5146	0.5214	0.5158
ATK2	0.5586	0.5596	0.5627	0.5651	0.5662	0.5752	0.5671

Table 4.2: Root-mean-square error for three models and various magnitude ranges

From Table 4.2 it can be judged that the BMR2 model is overall the best model, having the smallest RMSE for the total dataset. Also in each magnitude class it is better than the ATK2 model. For the magnitude range for which the Bommer model was developed, it is marginally better than the Bommer model. Hence, BMR2 is our model of choice which we will use in the remainder of the report.

Figure 4.10 shows again the distance and magnitude dependent misfit of the Bommer model, but now added to it the misfit of the BMR2 model. From this figure it can be judged at which distance ranges the improvement is made and whether the model under- or overpredicts the data. The main improvement is within the first 2 km, where the overprediction of the Bommer model is largely removed. Also the overprediction between 6 and 12 km is reduced. The lower panel shows the misfit for the complete dataset. It can be seen that for most distances the mean misfit is negative. That is, also BMR2 somewhat overpredicts the data, especially for distances larger than 15 km, for which no model update was provided.

Without adapting the Bommer model, the recorded PGV at short range are overestimated (Figure 4.7). With the BMR2 model, this overprediction is largely mended. Figure 4.8(b) shows that the new model on average slightly overpredicts the recorded values. All the extreme values in the database (maximum PGV values recorded per event within 2 km distance) are captured comfortably within the P99 model, with the most extreme values just crossing the P90 line.

From Figure 4.8(b) it can also be read what minimum magnitude values are needed to reach the 2 mm/s threshold. For P50, P90 and P99, these are 2.18, 1.82 and 1.53, respectively.
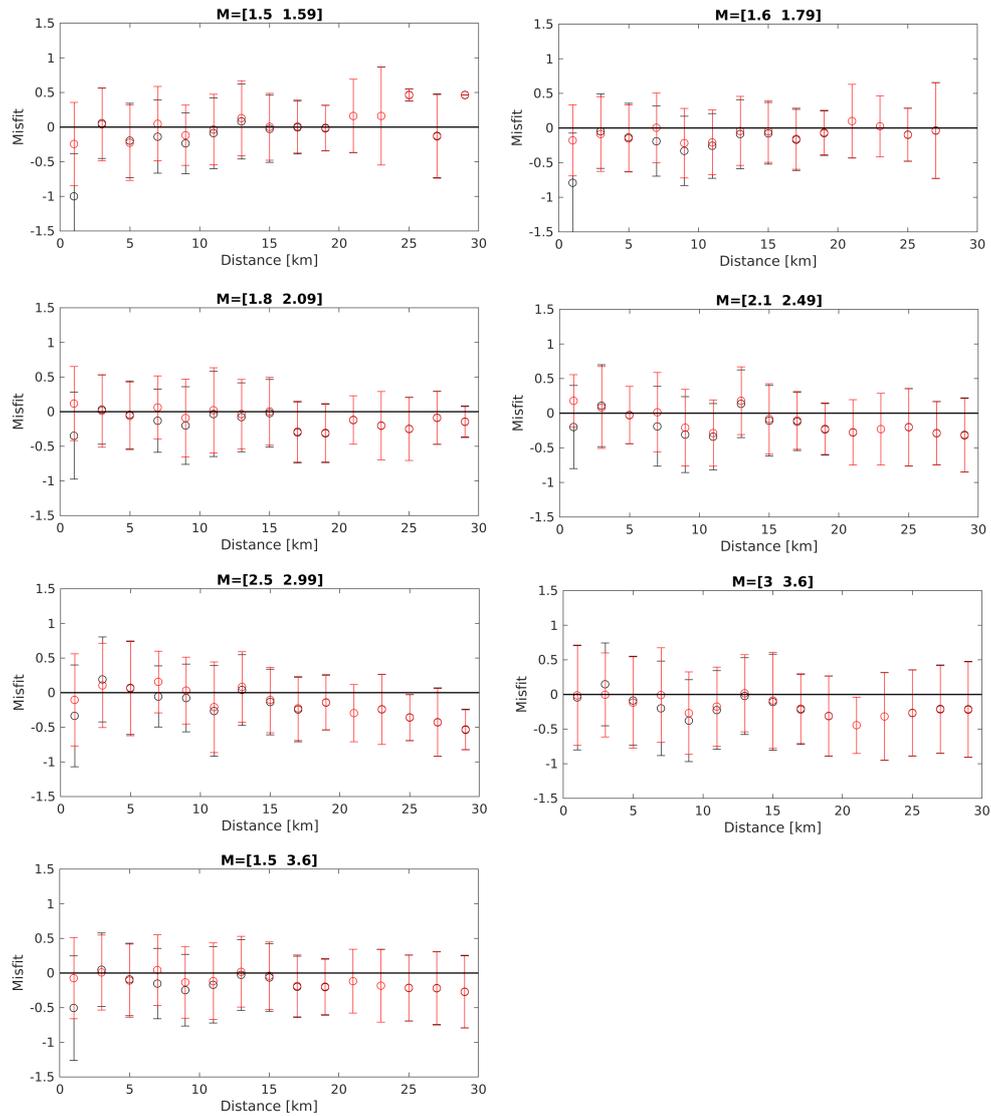
Figure 4.10: Misfit between PGVrot database and Bommer GMPE (black error bars) and the BMR2 GMPE (red error bars) for different magnitude ranges. The mean misfit (circles) and standard deviation (the bars denote from -1 to +1 standard deviation) are displayed over distance bins of 2 km. A few bins have unsufficient data to compute a meaningfull misfit distribution. This is, e.g.. the case for M=[1.5 1.59] events at large distances.

5 Combination of recordings and GMPE

In this chapter we develop a workflow for combining recordings and the adapted Bommer GMPE. We largely follow the approach taken by *Worden et al.* (2010) as is widely used at different seismological institutes around the world. The combination of model and records is done in two ways. On the one hand, the recordings yield information of average characteristics of the event and how these are different from the average GMPE. This difference is called event term and is discussed in the first section. On the other hand, local recordings are local 'ground truths' and can replace the model at the recording site. At distances away from the sensor, a weighted average is taken between the model and the recording. The weighing is controlled by the distance-standard-deviation model: a model that describes how the variability between samples increases as function of distance. In the second section, this model is determined for the PGV database. In the last section, the weights are implemented and discussed.

5.1 Event term

The event term expresses the average PGV level for which an individual event is different from the mean model. E.g., when the same amount of energy is released in a shorter duration (i.e., higher stress drop), the source-time function is more spiky and larger amplitudes will be recorded at the Earth's surface than for an event with the same magnitude, but with a longer-duration rupture. The way the event term is estimated, it may not be just a characteristic of the source. If the GMPE is applied in a setting in which the site amplification is —on average— larger or smaller than in the current PGV database, there is a so-called region-term. This region term manifests itself as an apparent event term.

When sufficient data are available, the (apparent) event term can be estimated from the recordings. The event term is computed as the L1-norm misfit of the recordings with the mean adapted Bommer model:

$$\epsilon = \frac{1}{n} \sum_{i=1}^{n} ln(Y_{obs,i}) - ln(Y_{mod,i}), \qquad (5.1)$$

where ϵ is the event term and n the number of recordings. This misfit is only computed when there are at least 3 recordings. The full event term is incorporated when 7 recordings are available. Hence, the event-term adjusted model is written as

$$Y_{mod}^{\epsilon} = Y_{mod} \exp(\frac{n}{7}\epsilon) \quad 3 \leq n < 7$$
$$Y_{mod}^{\epsilon} = Y_{mod} \exp(\epsilon) \quad n \geq 7. \qquad (5.2)$$

When an event term can be (fully) estimated, the between-event variability τ can be omitted. For a number of recordings n smaller than 7, τ will remain to have a finite size and needs to be included for computing P90 and P99 levels:

$$
\begin{aligned}
\tau^\epsilon &= \frac{7-n}{7}\tau \quad 3 \leq n < 7 \\
\tau^\epsilon &= 0 \quad n \geq 7.
\end{aligned}
\tag{5.3}
$$

Figure 4.10 shows that the misfit does not strongly vary with distance from the source. Thus, both nearby and more distance recordings can be used for estimating the event term.

5.2 Spatial correlation

Within Groningen, ground-motion maps are computed following largely the recipe from the USGS (*Wald et al.*, 2006). Following this reference, the resulting ground-motion maps are called shakemaps. A detailed near-surface velocity model has been derived in *Kruiver et al.* (2017), which forms the base for a site-amplification model (*Rodriguez-Marek et al.*, 2017) to which scenario-dependence is added in *Stafford et al.* (2017). This model is used to remove site effects from local recordings. Model and measured values are combined at the level of the first harder rock, the top of the Chalk Group (or equivalently, the base of the North Sea Group). The model and recordings are combined using an exponential model. The distance until which recordings are weighted in is largely determined the radius of influence r_{ROI}. The combined values are then propagated to the Earth's surface using the local transfer functions. Depending on the parameter of interest, this procedure leads to the combined PGV, PGA (peak ground acceleration) or SA (spectral acceleration) field.

For places outside Groningen, no detailed site-response model is known. As a consequence, the model and records need to be combined directly at the Earth's surface with the additional variability of the site response. This additional variability leads to a smaller r_{ROI} than for the Groningen shakemap implementation, which is based on a GMPE including site response (*Bommer et al.*, 2017). In this section we derive this r_{ROI} that includes the additional site-response variability.

The correlation lengths can be obtained from a covariance function or a semivariogram (*Schiappapietra and Douglas*, 2020; *Stafford et al.*, 2019). We take over the pragmatic approach as layed out in *Worden et al.* (2010) and *Boore et al.* (2003). In the latter approach, instead of a semivariogram, a distance-dependent standard deviation σ_{obs} is computed. This standard deviation can readily be compared with the within-event standard deviation ϕ from the GMPE. Doing so, the critical length scale r_{ROI}, at which $\sigma_{obs} = \phi$, is obtained.

We use the PGVrot database for magnitudes larger or equal than 1.8. To obtain a parameter that is approximately stationary (has similar average over an area) we first remove the modeled values from the observed values. Subsequently, we take the difference for all station pairs. E.g, for the first station pair between receiver 1 and 2 the difference is written as

$$
\Delta Y_1 = \{\ln(Y_{obs,1}) - \ln(Y_{mod,1})\} - \{\ln(Y_{obs,2}) - \ln(Y_{mod,2})\},
\tag{5.4}
$$

where Y is expressed in mm/s and Y_{mod} is generated with the BMR2 model. Thus, in fact the difference is taken of the (within-event) residuals. The above equation is repeated for all station combinations. Subsequently, this database of ΔY values is binned based on
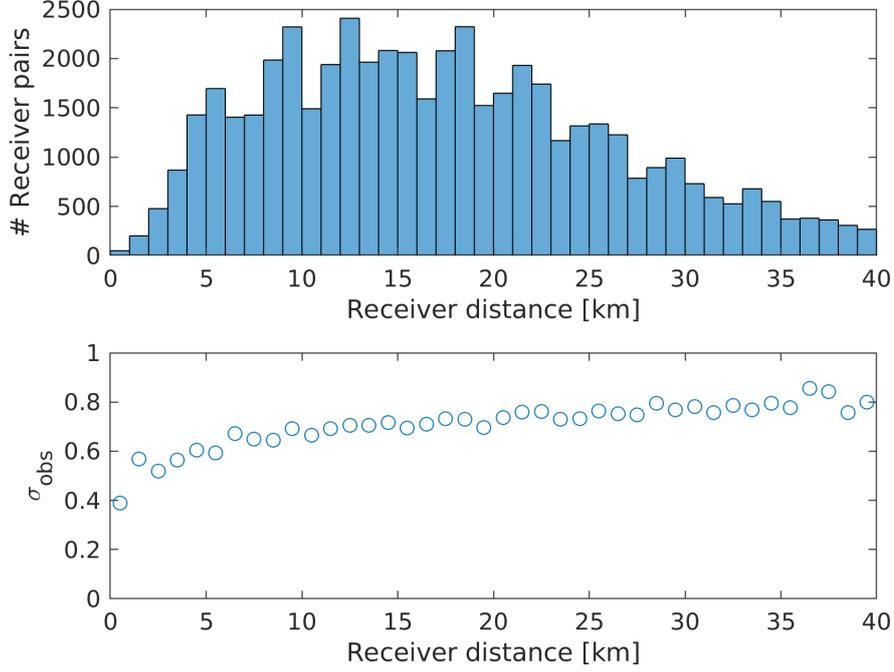
Figure 5.1: The distance standard deviation estimated with the PGV database expressed in PGVrot for $M \geq 1.8$. The upper panel shows the number of receiver pairs per 1 km bin. The lower panel shows the standard deviation of PGVrot residuals (equation 5.4) within each bin.

receiver-pair distance and for each bin i the distance standard deviation is computed:

$$\sigma_{obs,i} = \sqrt{\frac{1}{N_i} \sum_{j=1}^{N_i} (\Delta Y_j - \mu_i)^2}, \tag{5.5}$$

where N_i is the number of station pairs in a bin. The mean difference over residuals in a bin,

$$\mu_{obs,i} = \frac{1}{N_i} \sum_{j=1}^{N_i} \Delta Y_j, \tag{5.6}$$

is typically close to zero. When computing $\sigma_{obs,i}$ for all bins, using a binsize of 1 km, Figure 5.1 is obtained.

The experimental semivariogram would be determined with almost the same procedure (*Schiappapietra and Douglas*, 2020). Squaring equation 5.5 yields the (distance) variance instead of the (distance) standard deviation. Additionally, normalizing with $1/2N$ instead of $1/N$ yields the semivariance:

$$\gamma_{obs,i} = \frac{1}{2N_i} \sum_{j=1}^{N_i} (\Delta Y_j - \mu_i)^2. \tag{5.7}$$

In the following we will estimate a model that well describes the empirical distance standard deviation (Figure 5.1). We consider two different functional forms, fit unknowns with

least-squares and select the model with the smallest L2 norm of the residuals (i.e., sum of the squared residuals).

1) *Worden et al.* (2010) use the following model, which they took from *Boore et al.* (2003):

$$\sigma_{obs} = \sigma_{mod}(1 - exp(-\sqrt{0.6r_{obs}})), \tag{5.8}$$

where σ_{obs} is the distance standard deviation of the observation, σ_{mod} is the standard deviation of the model (GMPE) and r_{obs} is the distance between two observation points (receivers). Their equation is valid at distances smaller than r_{roi}. We use a small adaptation of their model to make it valid until large range:

$$\sigma_{obs} = \sigma_{obs,max}(1 - exp(-\sqrt{kr_{obs}})), \tag{5.9}$$

in which $\sigma_{obs,max}$ is the standard deviation for very large distances. This is an additional unknown that is estimated in the fitting procedure. This variability at large range will be somewhat larger than the model variability, which is a distance-independent average of residuals. The other unknown is the parameter k which controls how fast the variability rises as a function of distance.

2) The distance-standard-deviation model (equation 5.8) is a fit of datapoints that have been obtained primarily from an earthquake in California (*Boore et al.*, 2003) and which was obtained by 'eyeballing'. Instead of using this specific function, one could consider the exponential function that is widely used in Kriging (*Davis and Sampson*, 1986):

$$\sigma_{obs} = \sigma_{obs,max}\sqrt{1 - exp(-r_{obs}/a)}. \tag{5.10}$$

This function has two unknowns $\sigma_{obs,max}$ and the range a, which are estimated in the fitting procedure.

In the ideal case, $\sigma_{obs,max}$ could be obtained from σ_{mod}. Combining equations 5.5 and 5.7 one can see that $\gamma_{obs} = \frac{1}{2}\sigma_{obs}^2$. In case the residual is truly stationary, that implies that the semivariance approximates the model variance for large distances. That implicates (taking either equation 5.9 or 5.10 for r_{obs} going to infinity) that $\sigma_{obs,max} = \sqrt{2}\sigma_{mod}$. Since we have not determined σ_{mod} for this specific dataset, nor have the certainty that the residual is fully stationary, we estimate $\sigma_{obs,max}$ in the fitting.

To the models a nugget is added. This nugget represents small-scale variability within the recording site. The size of the nugget could directly be estimated from empirical data points. However, for our dataset there is a lack of data points at short distances (Figure 5.1) to do such estimation with confidence. We do not want to impose a large nugget as this would result in down-weighing the contribution of the recordings. As a compromise we take a nugget that is 20% of the within-event variability. For PGVrot (Table 3.2) this yields $0.2\phi = 0.10$.

Figure 5.2 shows the fit of the σ_{obs} data points with the two different models. The fit is made over data points up till 40 km distance. The fitted datapoints are only shown over the first 20 km, since the second 20 km is less eventful (Figure 5.1). The addition of the nugget and the least-squares fitting of the parameters k and a and $\sigma_{obs,max}$ leads to the following models:

1) adapted Boore model:

$$\sigma_{obs} = 0.1 + 0.691(1 - exp(-\sqrt{0.374r_{obs}})), \tag{5.11}$$

2) adapted exponential model:

$$\sigma_{obs} = 0.1 + 0.660\sqrt{1 - exp(-r_{obs}/5.30)}. \tag{5.12}$$
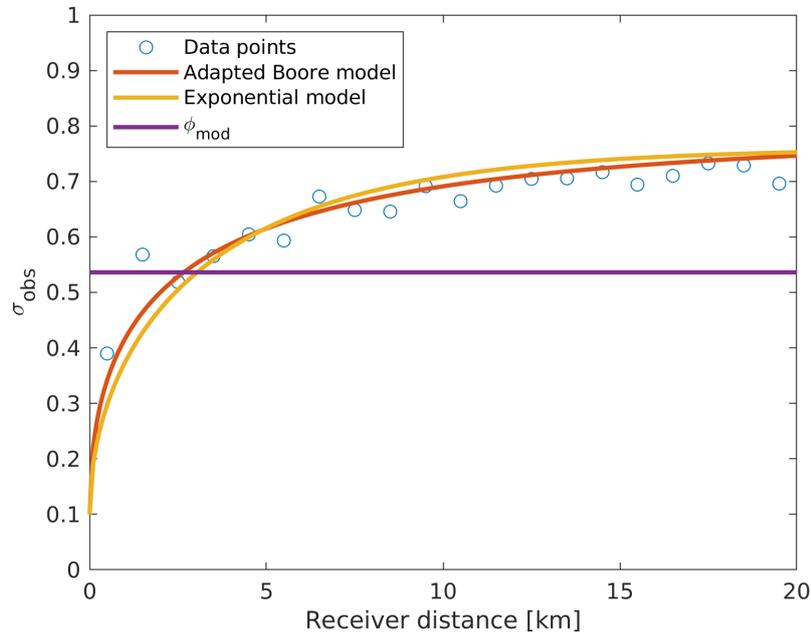
Figure 5.2: The same σ_{obs} data points as in Figure 5.1 over the first 20 km distance, with additionally 2 different model fits and the the within-event variability ϕ_{mod} of the Bommer model added.

The adapted Boore model and the adapted exponential model have a L2 norm of 0.0375 and 0.0684, respectively. This makes the adapted Boore model (equation 5.11) the preferred model.

In Figure 5.2 the model variability is added to obtain insights in the radius of influence. We assume that, in case there are data available to locally adapt the model, there are also sufficient data for estimating an event term. In that case, the model has ϕ_{mod} variability. For the adapted Boore model, the ϕ_{mod} level is crossed at 2.7 km. For that reason we will use $r_{roi} = 2.7$ km in the following.

The upper analysis could be repeated for PGVmax and PGVgeo. We do not expect that this will yield marked differences in distance-dependant standard deviation or estimated r_{roi}. Hence, we leave out this exercise.

5.3 Local perturbation

For combining model and measured values we take a similar approach as in *Worden et al.* (2010). For a specific grid point, the contribution of the nearby recordings and the model is dictated by the distance dependant standard deviation $\sigma_{obs}(r_{obs})$ and the distance independent variability of the model. For σ_{obs} the adapted Boore model is used (equation 5.11) from a receiver location until the radius of influence r_{roi}, which is the distance at which the recording has the same variance as the model and therewith an equal contribution to the combined PGV value. That is, both have a weight of 0.5 at this distance.

Furthermore, *Worden et al.* (2010) define a maximum radius r_{max}, which is larger than r_{roi}. Between r_{roi} and r_{max}, σ_{obs} is increased to a very large value such that the contribution of the recording is tapered to zero, with the following equation

$$\sigma_{obs} = \sigma_{roi} \frac{r_{max} - r_{roi}}{r_{max} - r_{obs}}, \tag{5.13}$$

where σ_{roi} is the standard deviation that is obtained with equation 5.11 at $r_{obs} = r_{roi}$. In the previous section we have made the choice such that $\sigma_{roi} = \phi_{mod}$. In the following we use $r_{max} = 4.0$ km

Following *Worden et al.* (2010), the variance of the combined PGV estimate becomes

$$\sigma_Y^2 = \frac{1}{\frac{1}{\sigma_{mod}^2} + \sum_{i=1}^{n} \frac{1}{\sigma_{obs,i}^2}}, \tag{5.14}$$

where the contribution from n recordings are weighted in. Since $\sigma_{obs,i}$ vary spatially, also σ_Y is a spatially varying function. The corresponding PGV at any location on the surface can be obtained with

$$Y = \left(\frac{Y_{mod}}{\sigma_{mod}^2} + \sum_{i=1}^{n} \frac{Y_{per,i}}{\sigma_{obs,i}^2} \right) \sigma_Y^2, \tag{5.15}$$

where Y_{mod} is the modeled PGV with a GMPE and Y denotes the combined PGV estimate. The perturbation on the model $Y_{per,i}$ is written as

$$Y_{per,i} = Y_{mod} \frac{Y_{obs}(\mathbf{x}_i)}{Y_{mod}(\mathbf{x}_i)}, \tag{5.16}$$

where $Y_{obs}(\mathbf{x}_i)$ is the recorded PGV at location $\mathbf{x}_i$ and $Y_{mod}(\mathbf{x}_i)$ is the modelled value for that location.

Figure 5.3 shows an example of combining the BMR2 model and one recording with equations 5.11 – 5.16. The example is 1D, showing the perturbation of the model in an azimuth in which there is a recording. This recording is at 5 km epicentral distance and has been chosen equal to the model. Consequently, P50 is not adapted. However, P90 and P99 are largely adapted due to the reduced variance of the combined PGV in a region around the recording. It is assumed that the model variability is described only by the within-event variability ϕ.

Figure 5.4 shows 2D examples of combining model and recordings. The first row shows an example with one station, which recording is exactly the value that is predicted by the mean GMPE. As a consequence, the P50 model (first column) is not adapted. The P90 and P99 models (second and third column, respectively) are perturbed around the station. This is because the total variance (equation 5.14) near the station is less than the model variance. Thus, the P90 and P99 levels are reached by a smaller increase of the mean GMPE.

The second row shows an example with one station, which recording is only half the amount that is predicted by the GMPE. As a result, the contours are becoming tighter in the azimuth where the station lies.

The third row shows an example with multiple stations, which either have double the predicted PGV (indicated by black triangles) or half the predicted PGV (red triangle). As was the case for the radiation pattern example (Figure 2.5) the maximum PGV level is shifted away from the epicenter. For P50 the contours are more extended in the north and south directions where the recordings with higher values lie. For P99, on the other hand, the contours are more extended to the west and east directions. This is because there are no stations east and west from the epicenter to reduce the total variance and hence a factor (larger than 2) needs to be added to the mean model to obtain P99 levels.
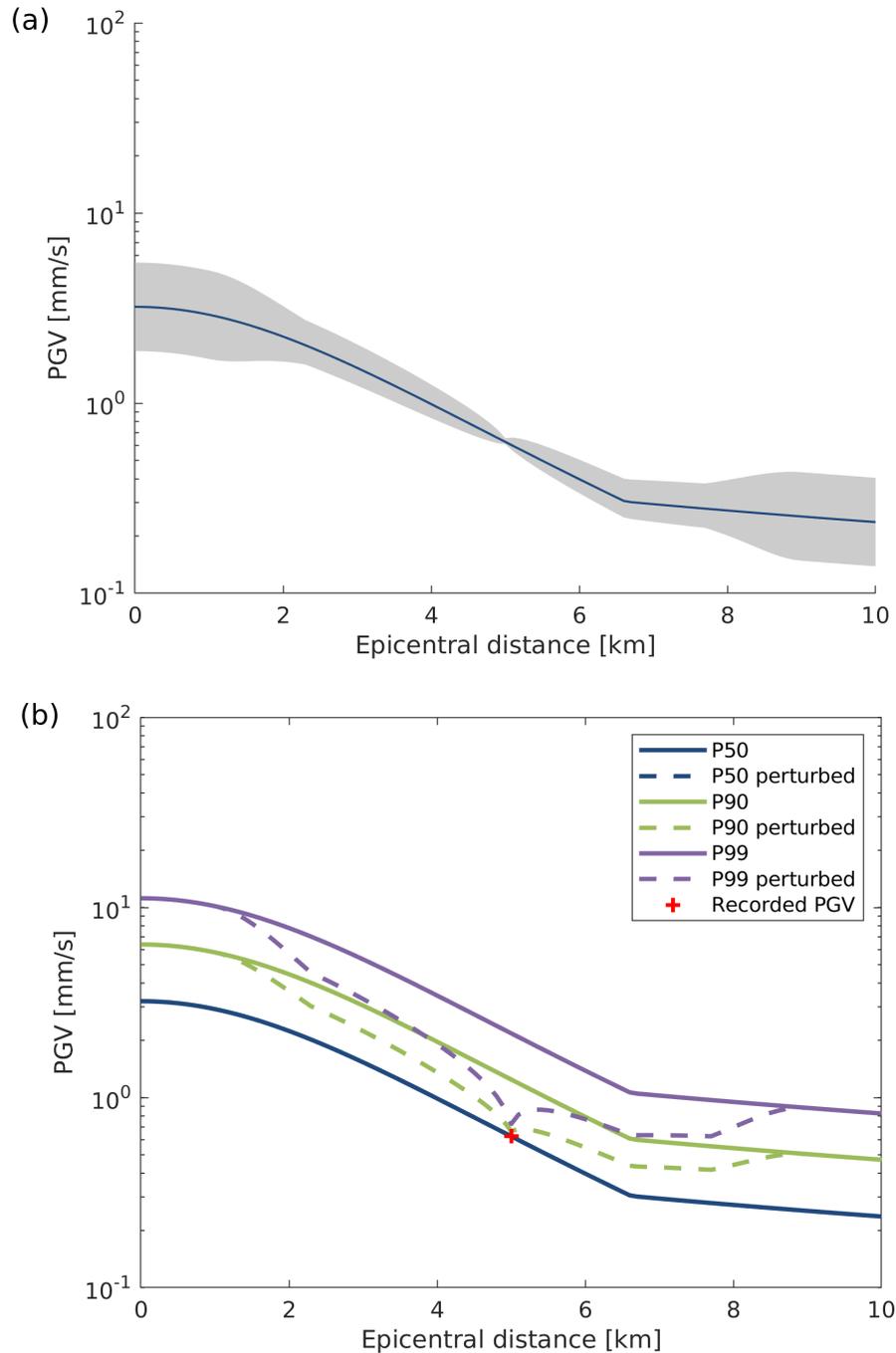
Figure 5.3: Synthetic 1D example of combining the BMR2 GMPE for M=2.4 and one local recording. The upper panel shows the mean (blue line) and 68 % confidence region (-1 to +1 σ_Y as defined in equation 5.14) for the perturbed model. The lower panel shows three different probability levels: P50, P90 and P99, both for the perturbed and unperturbed model. The recording at a distance of 5 km is indicated with a red cross.
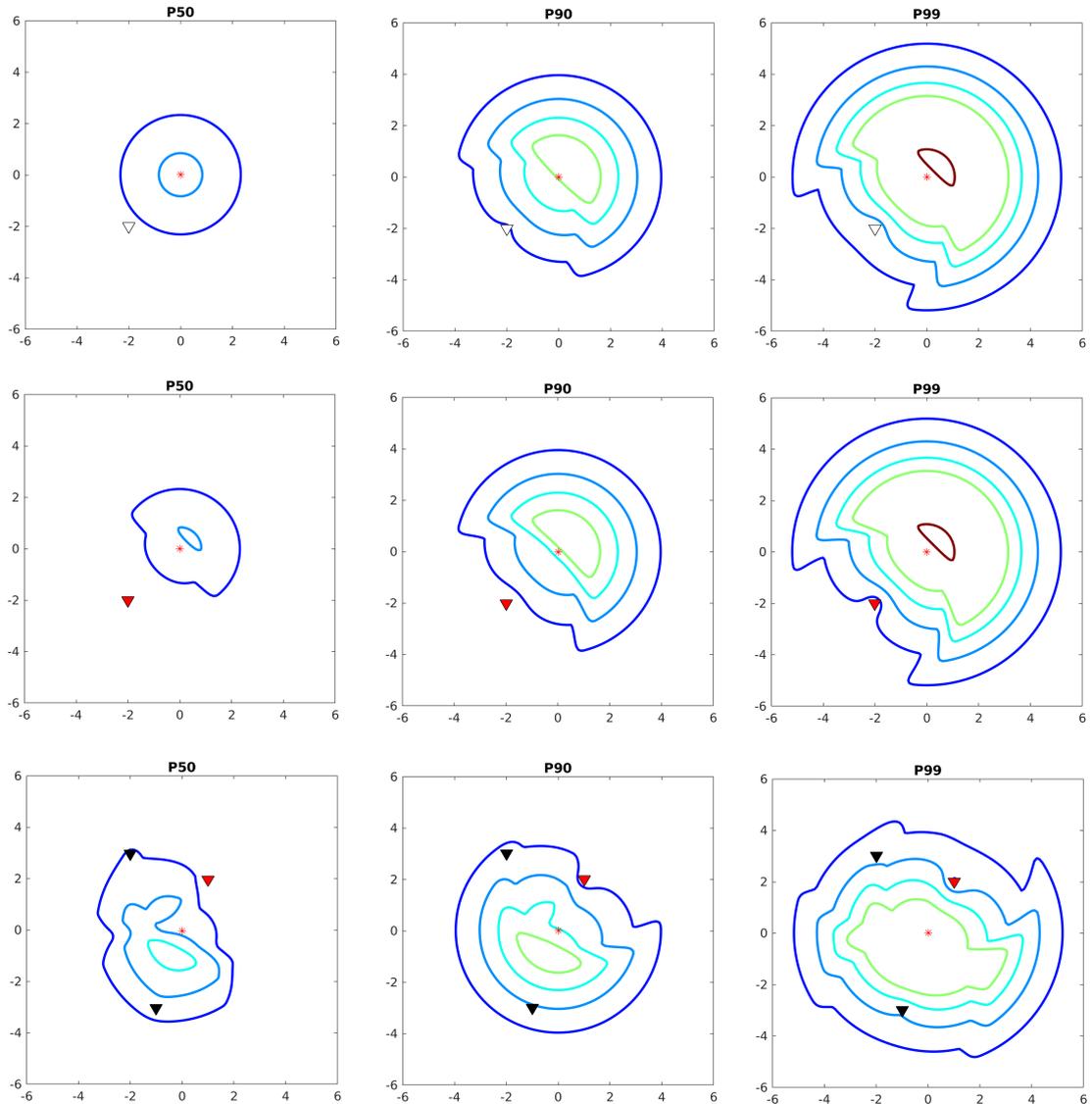
Figure 5.4: Examples of combining a GMPE and local recordings, for three different cases (rows) and three different probability levels (columns). The first column shows 50% of exceedance (P50), the second column 10% (P90) and the third column 1% (P99). The implemented GMPE is BMR2 for $M = 2.4$, with an intra-event variability. The stations (triangles) have either half (red), the same (white) or double (black) the PGV level as predicted by the GMPE. The contours are drawn at 2 mm/s (blue), 3 mm/s (light blue), 4 mm/s (cyan), 5 mm/s (light green) and 10 mm/s (red).

6 Proposed workflow

In this chapter we summarize the workflow for computing PGV threshold regions, as developed in the last chapters. This workflow is illustrated with three events outside Groningen.

The lowest relevant PGV threshold is 2 mm/s. In the current PGV database, this value is reached for a minimum magnitude of 1.9 (Section 2.2). The P99 from the BMR2 model (Figure 4.8b) passes this threshold level for M=1.53. For smaller magnitudes, even the P99 model does not reach the 2 mm/s threshold and hence a model-driven PGV threshold region would be non-existent. Threshold regions will be determined at the request of the 'Commissie Mijnbouwschade'.

In summary, we suggest the following workflow

1. Determine the epicenter and magnitude. A good distribution of stations nearby the epicenter will lead to a spatially confined epicentral probability density function.

2. Estimate the depth of the event. This depth can only be estimated from the seismic data if a few nearby stations pick up the event. If not, the depth is set by default to the reservoir depth level (gas-water contact in case of a gas reservoir). The connection with a reservoir can be made if the epicenter falls on top, or very close to a reservoir. If different reservoirs are stacked, knowledge on faulting and production can help in pinpointing the most likely seismogenic reservoir. If the event depth is considerably deeper than the reservoir depth, an anthropogenic origin of the event might be excluded.

3. If 3 or more PGV recordings have a SNR≥6 dB, start up the event term computation (Section 5.1). The event term is fully included from 7 PGV recordings onwards. With this number of recordings also the inter-event variability τ can be fully omitted and only the intra-event variability ϕ needs to be used.

4. Use the BRM2 GMPE (Section 4.4.1), with a further adaption of the event term, to compute a 2 mm/s PGV threshold region. If there are recordings within this region, or if there is a recording with ≥1 mm/s, then compute local perturbations of the modeled PGV field. The combined PGV map is computed following the recipe in Section 5.3.

5. Extract P50 PGV contours for 2, 3, 4, 5, 10 until maximum value with steps of 5 mm/s from the PGV field. Use the remaining variability to compute the 90th percentile (P90) and 99th percentile (P99) PGV field and extract the contours.

6. Publish a largely standardized report on the event with in it details on location, magnitude computation, PGV recordings and computation of PGV contours. The actual contours are distributed as an attachement in the form of kml files, which can be imported in, e.g., Google Earth, ArcGIS and QGIS. This report is made public within two weeks after the request.

The above workflow was largely implemented for the Zuidlaren 23-12-2016 event (*Ruigrok and Dost*, 2020). For that event, a large number of PGV recordings was available, which

made it possible to estimate an event term and to fully correct for this event term. However, the nearest ground-motion recording was at about 10 km epicentral distance, with PGV levels far below 1 mm/s. Thus, the relevant PGV contours (2 mm/s and up) were not perturbed by the local recordings.

Epicentral uncertainty is currently not being further propagated. The reason is that the PGV database and the empirical GMPE based on it (BMR2) already implicitly contain the average epicentral uncertainty.

In the following all figures are made using PGVrot. However, also the contours in terms of PGVmax and PGVgeo can be requested at KNMI as kml files.

6.1   Dalen 17-7-2018

6.1.1   Location

The Dalen event on 17-07-2018:9:53:57 with a local magnitude of 2.00 was detected by the KNMI network (*KNMI*, 1993) and located near real time with the Hypocenter method (*Lienert et al.*, 1986). This fast solution uses an average 1D model for the north of the Netherlands (*Kraaijpoel and Dost*, 2013). Here the epicenter is improved by using a best-fitting traveltime versus distance model based on a database of local P-wave traveltime picks. This data-driven model naturally incorporates actual underburden velocities and only well pickable phase arrivals. An error estimate is derived from the spread in picking times from the best-fitting model. This error incorporates both the local variations of the velocity field as well as picking errors. These errors are propagated further into an epicentral probability density function (PDF). This results into an updated epicenter and its 95% confidence region. For post-2014 events within the Groningen network, also an updated location is available (*Spetzler et al.*, 2018).

Fig. 6.1 shows the seismic sensors where manual P-wave picks are available for this event. A grid search is done for a region around the Hypocenter solution, as indicated by the red box in Fig. 6.1. In the first step, equal differential time (EDT, *Zhou*, 1994) residuals are computed. That is, for each grid point and for each station combination, the traveltime differences is forward modelled and tabulated. From these values, the observed traveltime differences are subtracted to obtain the EDT residuals. In the second step, the PDF is derived from the EDT residuals, using a L1 norm (*Tarantola*, 2005). Fig. 6.2 shows the 95% confidence area of the resulting PDF[1]. The location with the maximum probability is assigned to be the new epicenter. There is insufficient resolution to obtain a precise depth of the event. It is fixed to 3 km.

In the following the relevant parameters are listed. The new epicenter is listed both in wgs84 and in the Dutch national triangulation system (RD). Also a gridded version of the 95% confidence contour of the PDF, and its major and minor axes, can be found.

Epicenter in wgs84 [deg]: 6.714, 52.708

Epicenter in RD [m]: 244680, 525340

PDF major axis [m]: 4447

PDF minor axis [m]: 1784

Orientation of the PDF ellipse [deg]: -29.6

---

[1]Note that a smaller PDF would be obtained if also S-wave time differences and P-S delay times were included. For this purpose, the S-wave velocity model needs to be calibrated. This is work in progress.
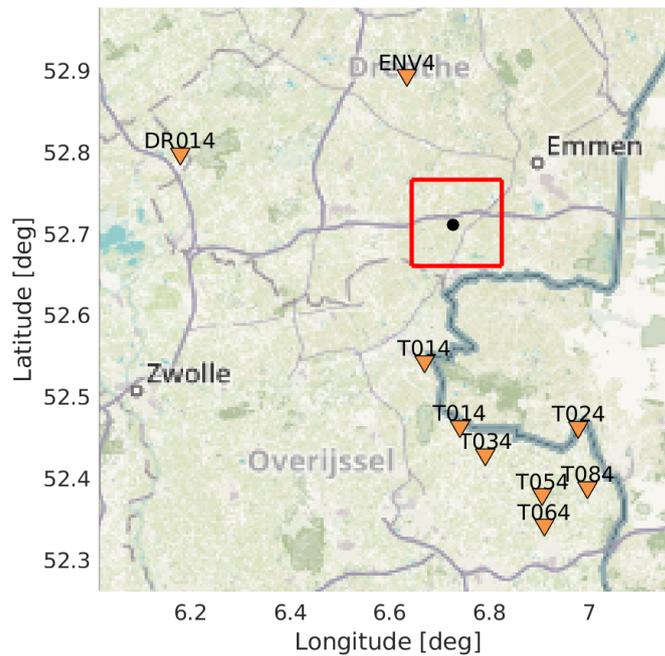
Figure 6.1: Overview map with locations of stations (orange triangles) where P-wave onsets were picked, the fast Hypocenter solution (black dot) and the boundary line of the area in which a grid search is done (red box). Background map is from www.openstreetmap.org.
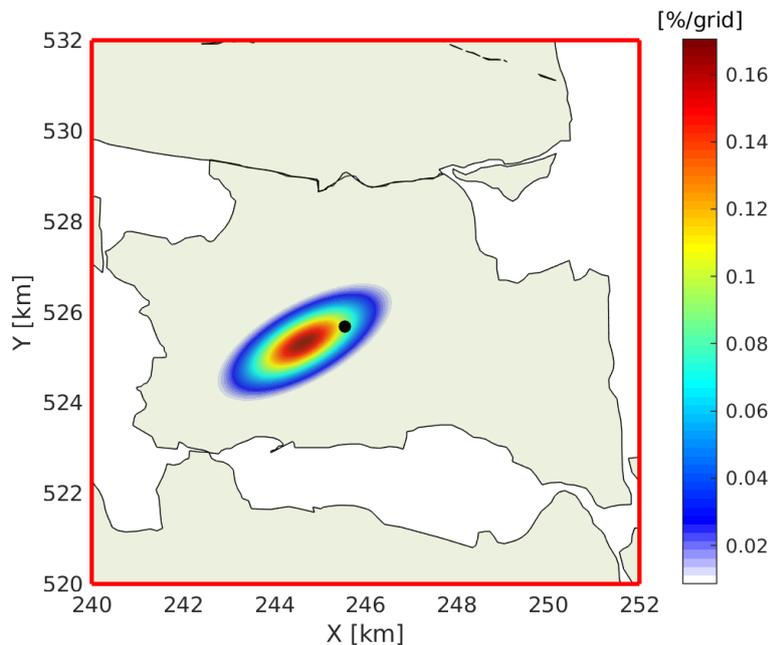
Figure 6.2: Map showing hydrocarbon fields (green-filled polygons), the fast Hypocenter solution (black dot) and the epicentral probability density function (PDF) using time-differences and an optimized model. The 95% confidence area of the PDF is shown, with probabilities expressed in percentage per grid point. The field polygons are from www.nlog.nl, using the March 2020 update.

95% confidence contour RDx [m]: 243050, 243240, 243480, 243720, 243900, 244080,
244260, 244440, 244620, 244784, 244920, 245100, 245225, 245400, 245520, 245671,
245820, 245940, 246060, 246180, 246300, 246407, 246488, 246584, 246642, 246636,
246540, 246408, 246180, 246000, 245820, 245580, 245400, 245220, 245040, 244860,
244680, 244500, 244373, 244200, 244054, 243900, 243779, 243614, 243480, 243360,
243240, 243120, 243000, 242895, 242820, 242734, 242690, 242702, 242813, 243050

95% confidence contour RDy [m]: 524020, 523987, 523987, 524019, 524055, 524103,
524160, 524227, 524303, 524380, 524448, 524545, 524620, 524729, 524811, 524920,
525038, 525139, 525249, 525370, 525504, 525640, 525760, 525940, 526120, 526360,
526513, 526600, 526660, 526671, 526660, 526626, 526585, 526533, 526474, 526406,
526330, 526244, 526180, 526087, 526000, 525904, 525820, 525700, 525595, 525492,
525381, 525258, 525120, 524980, 524860, 524680, 524500, 524320, 524140, 524020

The underlying waveform data used in the above analysis is publicly available and can
be obtained through

a GUI: http://rdsa.knmi.nl/dataportal/

FDSN webservices: http://rdsa.knmi.nl/fdsnws/dataselect/1/

6.1.2   PGV threshold regions

The Dalen event had a magnitude of 2.00. Since it is larger than 1.5, the PGV thresh-
old computation is started up. All accelerometer recordings at distances smaller than
$R_{max} = 6 + 40M = 86$ km are evaluated, which yields five recordings with an accept-
able SNR. Fig. 6.3 shows the PGV values as function of epicentral distance, together with
the event-term shifted BMR2 model for M=2.00 and an event depth of 3 km.

The event term is computed with equation 5.1 yielding a value of -0.334 and an effective value
of -0.239 (equation 5.2). The between-event variability τ is reduced to 2/7th of 0.252 (equa-
tion 5.3 with $n = 5$). The remaining model variability becomes $\sigma_{mod} = \sqrt{\phi^2 + \tau^2} = 0.541$.
This remaining variability is implemented to yield the confidence regions as plotted in Figure
6.3. The intersections of the P99 with the 2 mm/s and 3 mm/s thresholds yield a radii of
2.8 and 1.6 km, respectively. P90 crosses the 2 mm/s threshold at 0.9 km. P50 does not
reach any of the threshold values. Since there is no PGV recording within the 2 mm/s zone,
nor any PGV recording with a level larger than 1 mm/s, the event-term corrected model is
not further locally perturbed.
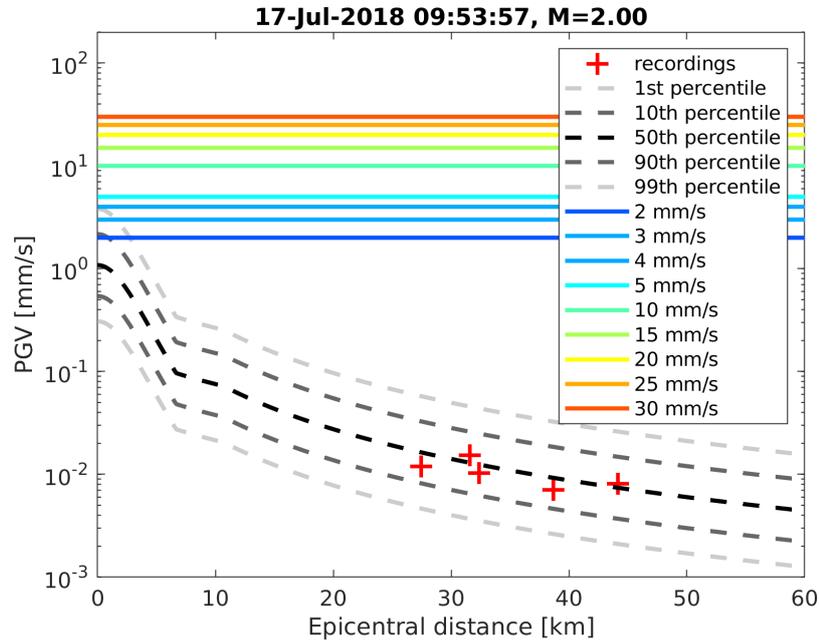
Figure 6.3: BMR2 GMPE expressed in PGVrot, confidence regions for that model (dashed lines), PGV thresholds (coloured lines) and measured PGVrot values for the Dalen event (red crosses).

The PGV contours for P90 and P99 are shown in Figures 6.4 and 6.5, respectively. [h]
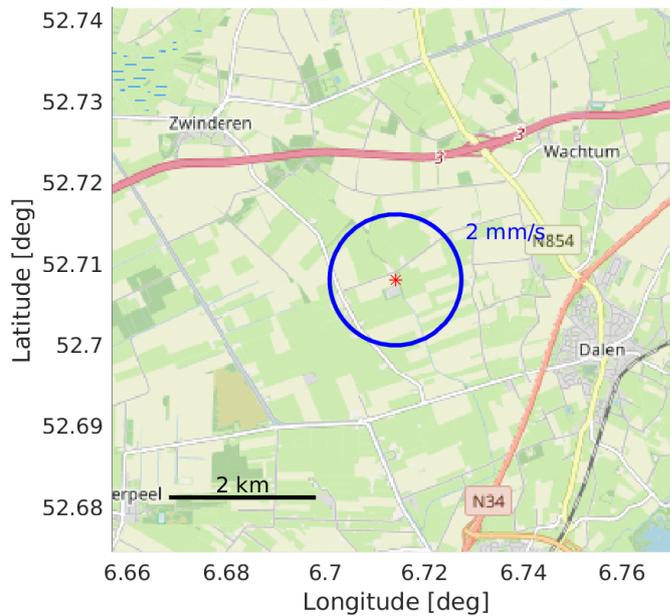
Figure 6.4: Circle depicting the bounding line of the 2 mm/s PGV threshold region for the P90 model, and the updated epicenter (red star).
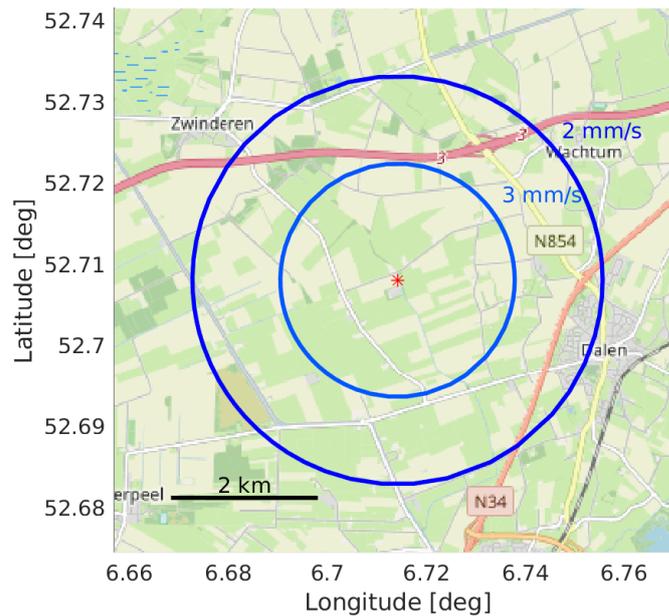
Figure 6.5: Circles depicting the bounding lines of the 2 mm/s and 3 mm/s PGV threshold regions for the P99 model, and the updated epicenter (red star).

6.2 Warder 04-06-2018

6.2.1 Location

The Warder event on 04-06-2018:23:01:02 with a local magnitude of 2.47 was detected by the KNMI network (*KNMI*, 1993) and located near-real time with the Hypocenter method (*Lienert et al.*, 1986). Subsequently, an updated epicenter and uncertainty were estimated with the same approach as described for the Dalen event. Fig. 6.6 shows the seismic sensors where manual P-wave picks are available for this event. A grid search is done for a region around the Hypocenter solution, as indicated by the red box in Fig. 6.6. Fig. 6.7 shows the 95% confidence area of the resulting PDF. The location with the maximum probability is assigned to be the new epicenter. There is insufficient resolution to obtain a precise depth of the event. It is fixed to 3 km.

In the following the relevant parameters are listed. The new epicenter is listed both in wgs84 and in the Dutch national triangulation system (RD). Also a gridded version of the 95% confidence contour of the PDF, and its major and minor axes, can be found.

Epicenter in wgs84 [deg]: 5.007, 52.549

Epicenter in RD [m]: 129200, 506900

PDF major axis [m]: 9778

PDF minor axis [m]: 4303

Orientation of the PDF ellipse [deg]: 13.6

95% confidence contour RDx [m]: 127562, 128000, 128469, 128800, 129100, 129352, 129600, 129800, 130020, 130215, 130400, 130594, 130732, 130892, 131000, 131124,
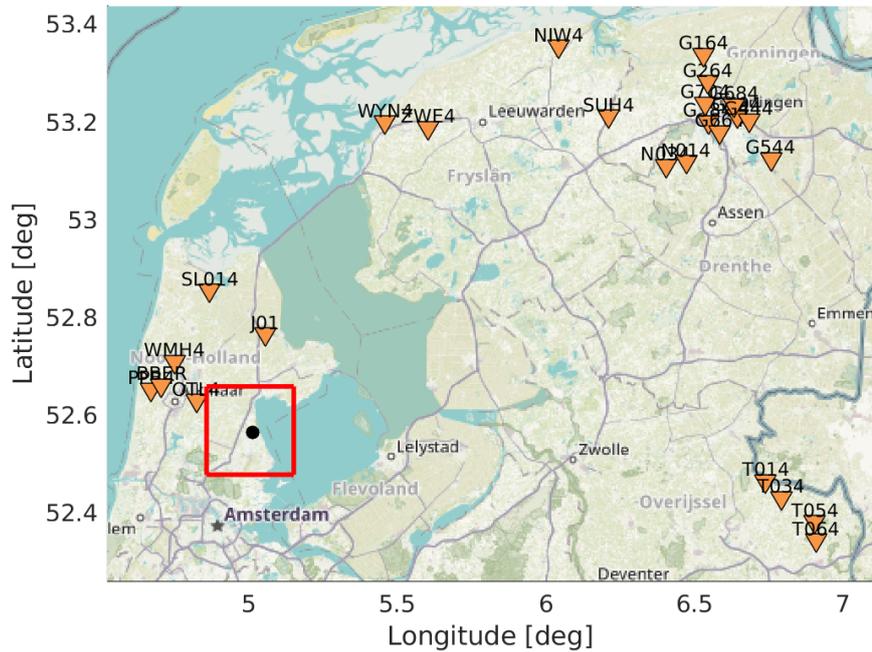
Figure 6.6: Overview map with locations of stations (orange triangles) where P-wave onsets were picked, the fast Hypocenter solution (black dot) and the boundary line of the area in which a grid search is done (red box). Background map is from www.openstreetmap.org.
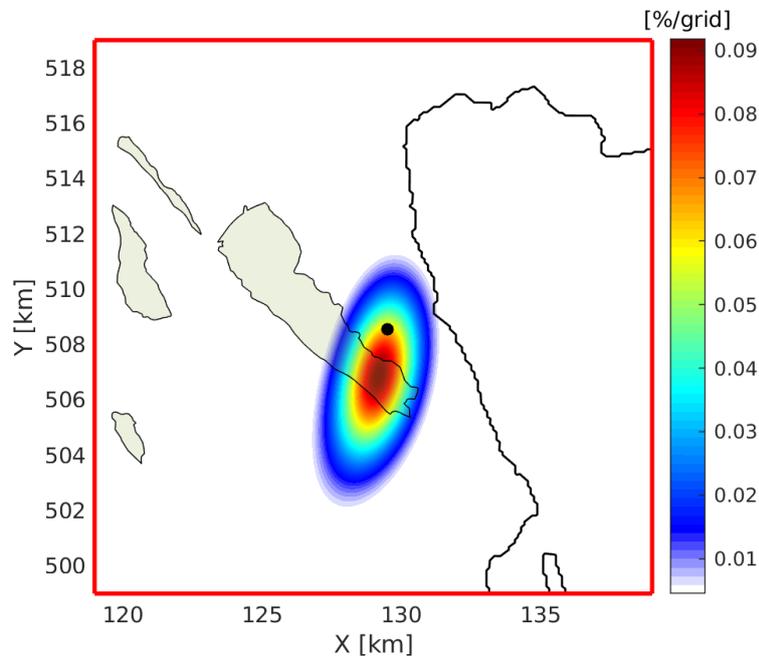
Figure 6.7: Map showing hydrocarbon fields (green-filled polygons), the fast Hypocenter solution (black dot) and the epicentral probability density function (PDF) using time-differences and an optimized model. The 95% confidence area of the PDF is shown, with probabilities expressed in percentage per grid point. The field polygons are from www.nlog.nl, using the March 2020 update.

131229, 131313, 131394, 131440, 131467, 131463, 131423, 131362, 131269, 131137, 131000, 130800, 130570, 130300, 129900, 129500, 129104, 128854, 128600, 128400, 128200, 128000, 127842, 127700, 127546, 127404, 127300, 127186, 127079, 126985, 126900, 126816, 126760, 126712, 126690, 126688, 126711, 126782, 126887, 127009, 127200, 127562

95% confidence contour RDy [m]: 502000, 501902, 502000, 502164, 502376, 502600, 502859, 503101, 503400, 503700, 504021, 504400, 504700, 505100, 505408, 505800, 506200, 506600, 507100, 507500, 508000, 508500, 509000, 509400, 509800, 510200, 510503, 510835, 511100, 511302, 511433, 511393, 511200, 511000, 510733, 510474, 510169, 509814, 509500, 509177, 508800, 508400, 508080, 507700, 507300, 506900, 506481, 506000, 505600, 505100, 504700, 504200, 503800, 503300, 502900, 502600, 502297, 502000

6.2.2 PGV threshold regions

The Warder event had a magnitude of 2.47. Since it is larger than 1.5, the PGV threshold computation is started up. All accelerometer recordings at distances smaller than $R_{max} = 6 + 40M = 106$ km are evaluated, which yields 1 recording with an acceptable SNR. Fig. 6.8 shows the PGV value at its epicentral distance, together with the BMR2 model for M=2.47 and an event depth of 3 km.

Since there is only one PGV recording, no event term can be estimated and the mean adapted Bommer model is used with the complete model variability $\sigma_{mod} = 0.59$. This variability is implemented to yield the confidence regions as plotted in Figure 6.8. The intersections of the P99 with the 2 mm/s threshold yields a radius of 5.9 km. For P90 and P50 this same threshold is reached at 4.5 and 2.8 km, respectively. Since there is no PGV recording within the P99 2 mm/s zone, nor any PGV recording with a level larger than 1 mm/s, the model is not further locally perturbed.
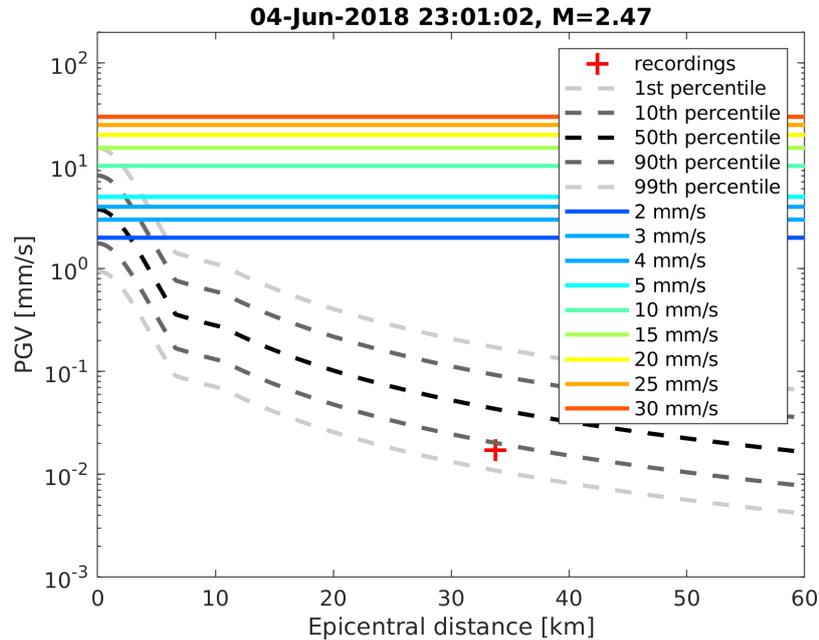
Figure 6.8: BMR2 GMPE expressed in PGVrot, confidence regions for that model (dashed lines), PGV thresholds (coloured lines) and one measured PGVrot value for the Warder event (red cross).

The updated epicenter together with the computed radii for the P50, P90 and P99 models (Figure 6.8) define the PGV threshold regions. On Figure 6.9 the bounding lines for P90 are shown. Similarly, Figures 6.10 and 6.11 show the threshold regions for the P90 and P99 models, respectively.

We repeat the upper analysis of computing the P50 2 mm/s radius for other definitions of PGV. In the upper analysis we used PGVrot and found a radius of 2.8 km. If we use PGVgeo instead, the model values are reduced by a factor of 0.6074 (Section 4.4.1) and the radius is (nonlinearly) reduced to 1.2 km. If we use PGVmax, the model values are reduced by a factor of 0.9218 and the radius is reduced to 2.6 km.
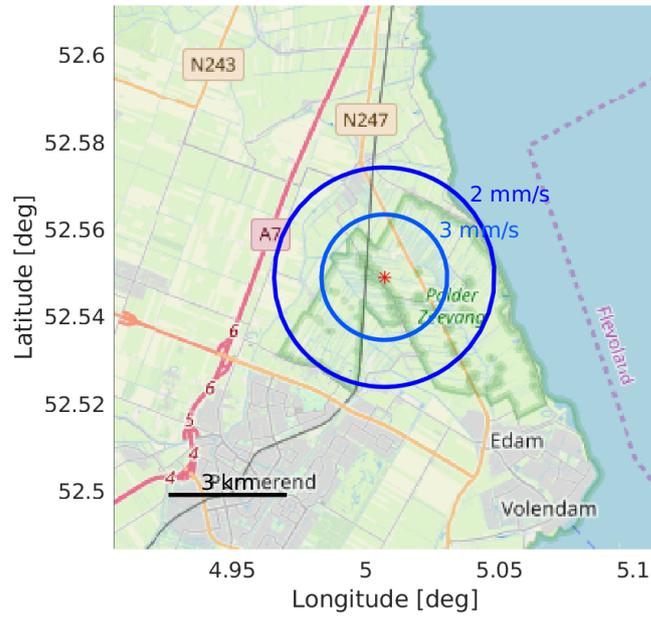
Figure 6.9: Circles showing the bounding lines of the 2 mm/s PGV threshold regions for the P50 model, and the updated epicenter (red star).
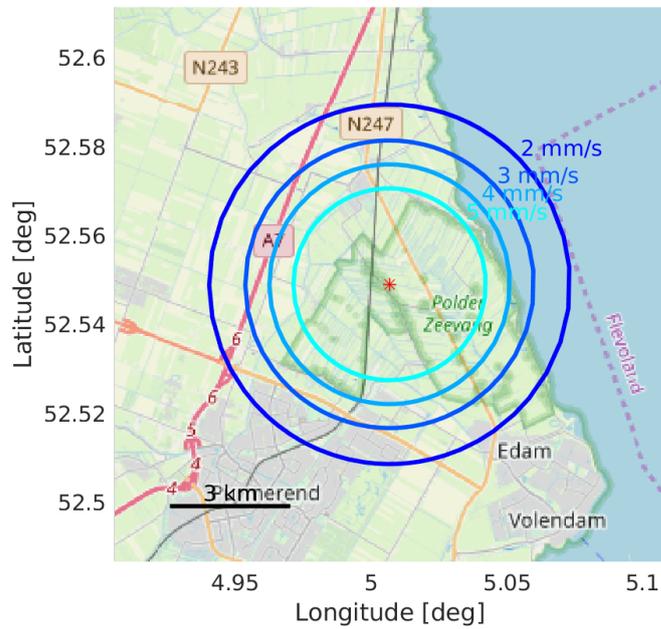
Figure 6.10: Circles showing the bounding lines of the 2, 3 4 and 5 mm/s PGV threshold regions for the P90 model, and the updated epicenter (red star).
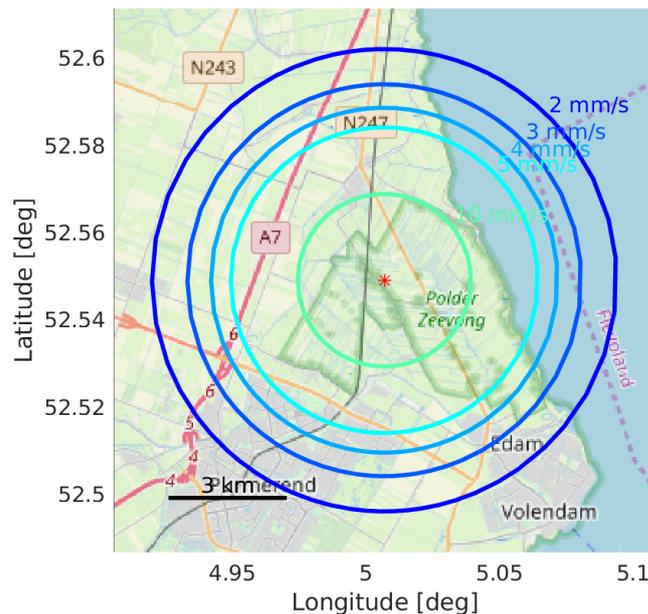
Figure 6.11: Circles showing the bounding lines of the 2, 3 4, 5 and 10 mm/s PGV threshold regions for the P99 model, and the updated epicenter (red star).

6.3   Roswinkel 25-10-2000

The events considered thus far did not have local recordings to perturb the modeled PGV field. For the Roswinkel events there was an accelerometer network that had been laid out directly above the gas field, resulting in local PGV recordings. From these Roswinkel events we select the one with $M \leq 3$ with the most PGVs in the database. This the Roswinkel 25-10-2000 event which had a magnitude of 3.20.

For this event we do not compute a new epicenter. Since the event was located, amongst others, with the 4 nearby stations, already a precise epicenter is available. We exclude the PGV obtained at station ROS3, since this station was located in a large building for which soil-structure interaction could not be excluded.

Fig. 6.12 shows the PGV values as function of epicentral distance, together with the event-term shifted BMR2. The effective event term (event term weighted by number of observations) is computed with equation 5.1 and 5.2 yielding a value of -0.057. The between-event variability τ is reduced to 3/7th of 0.252 (equation 5.3 with $n = 4$). The remaining model variability becomes $\sigma_{mod} = \sqrt{\phi^2 + \tau^2} = 0.547$. This remaining variability is implemented to yield the confidence regions as plotted in Figure 6.12.

The model values were computed for an event depth of 3 km, although the event was estimated to be nucleated at 2 km depth, at reservoir level. The reason is that for estimating the magnitude, an attenuation relation was used that assumes a depth of 3 km. Together with the usage of nearby recordings for magnitude estimation this results in the higher PGV levels being already incorporated in the magnitude and do not need to be counted twice by also reducing the depth of the hypocenter. It can be seen in Figure 6.12 that a model depth of 3 km results in well capturing the measured PGVs within the confidence region.
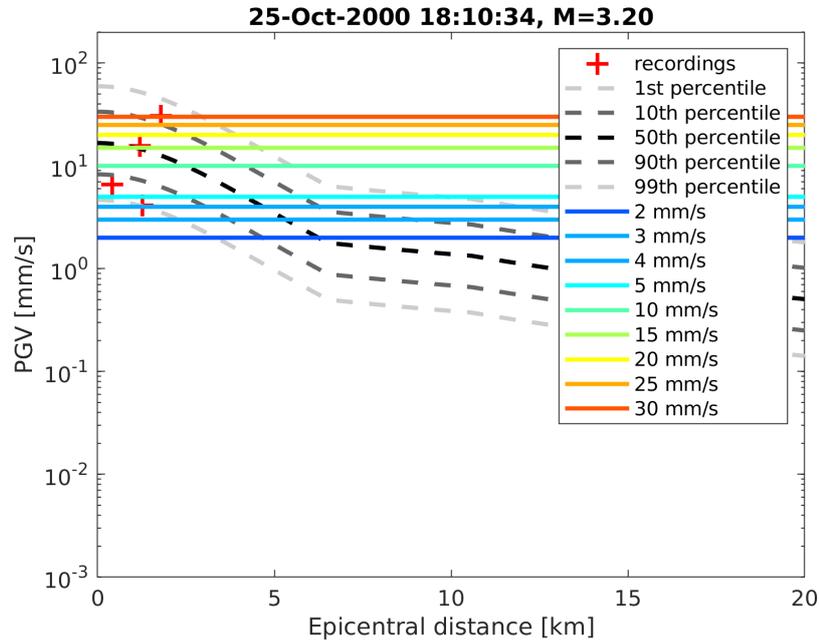
Figure 6.12: BMR2 GMPE expressed in PGVrot, confidence regions for that model (dashed lines), PGV thresholds (coloured lines) and measured PGVrot values for the Roswinkel event (red crosses). The lowest two PGVs have almost the same amplitude and distance and are hence plotted on top of each other.

The intersections of the P99 with the 2 mm/s threshold yields a radius of 18.9 km. All 4 stations are within this radius and are thus used to perturb the modeled PGV field using the approach as detailed in Section 5.3. Fig. 6.13 shows the resulting threshold regions for P50, P90 and P99. The PGV field shows a high lobe in NW of the epicenter which is likely related to normal faulting on a SW-NE striking fault (compare with the modeled radiation pattern shown on Fig. 2.5).
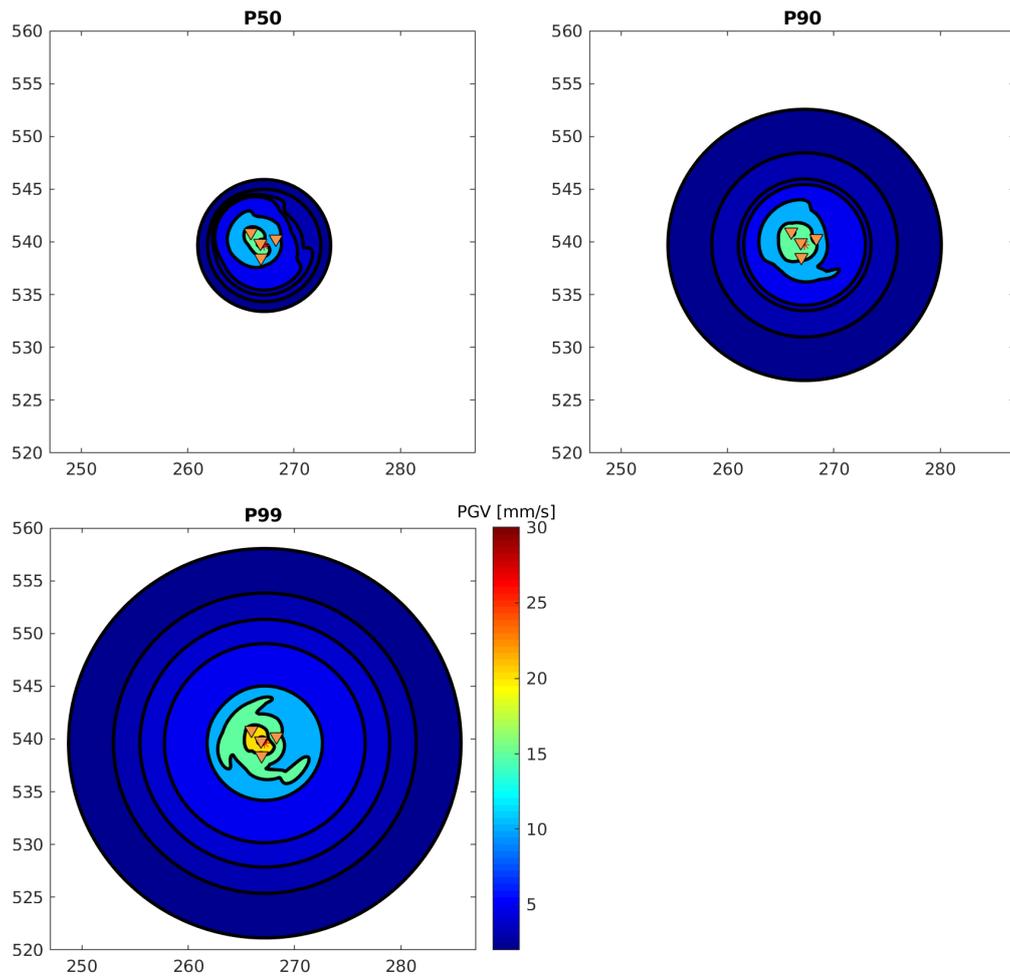
Figure 6.13: Filled contours showing the combined modeled and measured PGV field for the Roswinkel 3.2 event, for a 50% chance of exceeding (upper-left panel), a 10% chance of exceeding (upper-right panel) and a 1% chance of exceeding (lower-left panel). The red star denotes the epicenter, orange triangles denote seismic stations where a PGV value was obtained. The threshold regions are shown in a local coordinate system in kilometers (Amersfoort / RD New).

7 Discussion and Conclusions

7.1 Current work

In the preceding we developed a workflow for computing PGV threshold regions for events outside of Groningen. We illustrated the workflow with three events.

First we compiled a database of PGVs from past recordings in the northeast of the Netherlands. We evaluated 4 different GMPEs, all at least partly based on historic Dutch data. The Dost GMPE is based on an attenuation relation that has been shifted using near-range strong motion recordings. The attenuation beyond 5 km fits well with the PGV database, but the modest decay in amplitudes over the first 5 km does not fit well. The same holds for the Douglas GMPE which has similarly modest decay over the first 5 km. Moreover, at farther range, Douglas predicts a faster decay in amplitudes than what is observed in the PGV database. The Atkinson has stronger decay over the first 5 km, which better fits the data. The overall amplitude levels are, however, much lower than what is observed in the Netherlands. A contribution factor would be that the model was primarily derived with data from the Geysers region, which has smaller near-surface amplification than what is found in the Netherlands. The Bommer GMPE fits best with the data both for events within and outside Groningen. This is not surprising since the Bommer GMPE is the most recent model which has been obtained with a large subset of the PGV database that we have used. The Bommer GMPE has by far the richest empirical base, with a dense sampling of recordings over the first 35 km distance. It is likely that induced events outside Groningen are also similar normal faulting events at similar depth ranges and with no dramatically different near-surface settings.

In *Ruigrok and Dost* (2020) it was thought necessary to split up the workflow in two different geological settings: one for events with a high impedance contrast in the overburden and one without. In this report, however, we found that the (adapted) Bommer model is equally suited for induced events with and without salt cover. The small fraction of PGV values from a setting without salt cover (Roswinkel) is still well described with the generic GMPE. The underlying reason is likely that for both cases the same attenuation relation is used for computing the magnitude. Without the presence of a salt cover, the fraction of radiated seismic energy that reaches the Earth's surface is higher. This results in a higher magnitude estimation than for the setting with a sealing salt cover. This higher magnitude results in modeling higher PGV values, which well match the observed higher PGV values. In other words, the differences in PGV levels for different overburden settings are already incorporated in the magnitude estimation. The drawback of this approach is a worse match of magnitude with radiated energy. The advantage is a simplified workflow for PGV threshold computations in which no discrimination needs to be made for the specific geological setting.

We found that the Bommer model overestimates the PGV recordings over the first two kilometers epicentral distance, mainly for small magnitude events outside the range for which the GMPE was developed. Since especially this near range is important in terms

of possible damage, this prompted the derivation and testing of adapted models. We took both the Bommer and Atkinson functional forms as a base and re-calibrated coefficients with the total NL PGV database (M=[1.5 3.6]). The adapted models were coined BMR2 and ATK2, respectively. A distinctive feature in the Atkinson functional form is that there is a M^2 dependency of PGV. The fitting exercise showed that this term is very small. A distinctive feature of the Bommer functional form is that the geometrical spreading has been adapted to the Groningen situation, with three different distance ranges. Especially this last feature resulted in a better performance of the BMR2 model. This adapted Bommer model largely resolves the overprediction in the first 2 km epicentral distance and partly resolves the overestimation between 6 and 10 km. The attenuation at larger range is not adapted. BMR2 remains a conservative model, with on average a small overprediction of the observed amplitudes, for most distances and magnitudes.

With forward modeling the ground motion of a typical induced event we showed that the PGV levels fluctuate largely in the source area. The maximum PGV level is shifted away from the epicenter, with a distance dependent on the dip and the depth of the fault. For the Zeerijp 08-01-2018 event we found a distance of about 1 km. This strong fluctuation of ground motion in the source region, together with an inherent uncertainty of the epicenter, can explain why the mean amplitude levels over the first few kilometers epicentral distance should be flatter than in the original Bommer model.

The adapted model BMR2 has as an added benefit the rewriting of epicentral distance to hypocentral distance. This makes it possible to use the equation in settings where the hypocentral depth is different than in Groningen.

In our proposed workflow, the BMR2 model, together with local recordings, is used for computing PGV contours. We largely follow the approach set out by *Worden et al.* (2010). On the one hand, the local PGV recordings are used to estimate and remove an event term; a term that describes how an individual event is different from the average event in the database. On the other hand, the PGV recordings are used as ground truths at their recording site and away from that site they are weighted in together with the model to give a combined PGV field. The influence of the recordings on the PGV field is dictated by the radius of influence r_{roi}.

We added a few changes with respect to the workflow described in *Worden et al.* (2010). First we slightly adapted the distance-standard-deviation model into equation 5.9. This was to ensure that at r_{roi} there is an equal contribution from model and recording. With their model, there remains to be a larger contribution for the recording at this distance [Figure 2 in *Worden et al.* (2010)]. Secondly, we added a nugget of 20% of the within-event variability. Thirdly, we fitted the standard deviation model with the Dutch PGV database and found $r_{roi} = 2.7$ km. This is considerably smaller than the value of $r_{roi} = 10$ km which is used in *Worden et al.* (2010). The considerably smaller correlation lengths in the Dutch setting can have multiple reasons. The induced events are much shallower than the tectonic events in California, USA. Deeper events will result in smaller spatial variations in the PGV field. Furthermore, site conditions likely vary more drastically in the Netherlands and these site conditions were not removed prior to computing the distance standard deviation. These site conditions are also not removed when computing PGV threshold regions for events outside Groningen, until they are well quantified.

In the combined model, the PGV recordings have a large influence. First, they globally influence the model by estimating and implementing an event term. Second, they locally perturb the model. An important influence is the reduced combined variance in the region around the receivers. As an effect, smaller PGV increments need to be added to reach P90

and P99 levels. When stations record larger PGV levels than predicted by the mean model, still P90 and P99 levels can be lower than for a pure model-driven PGV field (Figure 5.4). Besides that, nearby receivers play an important role in reducing the uncertainty of the epicenter, in estimating the depth of the event and to establish a magnitude. In general, the existence of local measurements will result in less uncertainty and therefore in a smaller damage probability contour.

We propose to use the variability as found in *Bommer et al.* (2019) also for events outside Groningen. This variability is caused, among others, by variation in site condition and azimuthal variation of source radiation. In Groningen, there is a large variation in site conditions, varying from pleistocene sands with modest near-surface amplification to medium-deep holocene deposits with large amplification (*Van Ginkel et al.*, 2019). Most of the Netherlands is covered with either a pleistocene or pleistocene and holocene blanket. Hence, it is likely that outside of Groningen a similar variability is encountered. Also variability caused by source radiation is likely similar for Groningen and regions outside. Seismicity, both inside and outside Groningen, is characterized by normal rupture on steeply dipping faults.

We illustrated the workflow for three events that occurred outside Groningen. For the Dalen event, five PGVs could be derived from stations in the region to compute an event term. No local stations were available to perturb the modeled PGV field. For the Warder event there was a lack of both regional and local stations. As a result, we had to resort to the mean BMR2 model using the full variability. As a consequence, the P99 PGV threshold regions are relatively large for this event. For the Roswinkel event, on the other hand, there were 4 local stations, which resulted in a largely (local-data) driven PGV field in the source area. This field shows large complexity with multiple highs, which can be explained by the radiation pattern together with local site effects.

7.2 Future directions

A future improvement would be to incorporate site terms in the GMPE. These would first need to be estimated. At the accelerometer sites, a detailed site response model could, e.g., be obtained with seismic cone-penetration tests. At other locations, proxies of site response would need to be available, e.g., in the form of V_{S30} maps. If these become available, a site-term corrected GMPE can be derived and used, which would have lower variability.

Bommer et al. (2016) conducted a test with including a site-term proxy in their empirical-GMPE regression for Groningen. They included a V_{S30} variable in their GMPE and found only a very weak dependence of predicted ground motion on V_{S30}. In the meantime, however, near-surface velocity models have improved. Moreover, site-term proxies could be derived that are better suited to the specific Dutch-induced case with shallow, low-magnitude seismicity and very soft soil conditions.

In this study, we largely followed the approach by *Worden et al.* (2010) for combining recordings and a GMPE, which is largely analogous to Kriging with a trend (`http://usgs.github.io/shakemap/manual3_5/tg_philosophy.html`). Drawbacks of this approach include the ad-hoc function (equation 5.13) with which a contribution of a recording is tapered to zero at a maximum distance. In the meantime, more delicate ways of including the recordings have been published. For a future update we would consider the Baysian Networks approach as introduced by *Gehl et al.* (2017), or the conditional multivariate normal distribution approach (*Worden et al.*, 2018).

In this study, the number of events and recordings outside Groningen was limited. It is anticipated that the database with outside-Groningen events will significantly grow over the coming years, with the installation of new strong-motion networks. This updated database

would allow a re-calibration of the GMPE. Moreover, in the current work we choose to retain the model variability from the Bommer empirical model. Also these could be re-calibrated with the extended PGV database.

The faulting mechanism has a strong influence on the PGV distribution (Figure 2.5). Currently, only an average radiation pattern of the typical normal-faulting mechanisms is included in the GMPE, leading to a radial pattern. When it becomes feasible to readily estimate moment tensors for events outside Groningen, the strike, dip and rake of the faulting could be incorporated into the GMPE, which would result in a more precise modeling of the PGV field.

The PGV database, as assimilated in the study, lacks a distribution of receiver pairs at short range. The database is dominated by the G-network, which has a station spacing of 4 to 5 km. This precluded an empirical estimation of a nugget when calibrating a distance-standard-deviation model. In the future, this nugget could be estimated from events recorded over the NAM flexible array and or surveys that are planned to quantify soil-structure interaction.

7.3 Acknowledgements

Bibliography

Atkinson, G. M. (2015), Ground-motion prediction equation for small-to-moderate events at short hypocentral distances, with application to induced-seismicity hazards, *Bulletin of the Seismological Society of America*, *105*(2A), 981–992.

Atkinson, G. M., E. Yenier, N. Sharma, and V. Convertito (2016), Constraints on the near-distance saturation of ground-motion amplitudes for small-to-moderate induced earthquakes, *Bulletin of the Seismological Society of America*, *106*(5), 2104–2111.

Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010), ObsPy: A Python toolbox for seismology, *Seismological Research Letters*, *81*(3), 530–533.

Bommer, J. J., P. J. Stafford, and M. Ntinalexis (2016), Empirical ground-motion prediction equations for peak ground velocity from small-magnitude earthquakes in the Groningen field using multiple definitions of the horizontal component of motion, *NAM Study Report*.

Bommer, J. J., B. Dost, B. Edwards, P. P. Kruiver, M. Ntinalexis, A. Rodriguez-Marek, P. J. Stafford, and J. van Elk (2017), Developing a model for the prediction of ground motions due to earthquakes in the Groningen gas field, *Netherlands Journal of Geosciences*, *96*(5), s203–s213.

Bommer, J. J., P. J. Stafford, and M. Ntinalexis (2019), Updated emperical GMPEs for PGV from Groningen earthquakes — March 2019, *NAM Study Report*.

Boore, D. M., J. F. Gibbs, W. B. Joyner, J. C. Tinsley, and D. J. Ponti (2003), Estimated ground motion from the 1994 Northridge, California, earthquake at the site of the Interstate 10 and La Cienega Boulevard bridge collapse, West Los Angeles, California, *Bulletin of the Seismological Society of America*, *93*(6), 2737–2751.

Cauzzi, C., R. Sleeman, J. Clinton, J. D. Ballesta, O. Galanis, and P. Kästli (2016), Introducing the European rapid raw strong-motion database, *Seismological Research Letters*, *87*(4), 977–986.

Davis, J. C., and R. J. Sampson (1986), *Statistics and data analysis in geology*, vol. 646, Wiley New York et al.

Dost, B., T. Van Eck, and H. Haak (2004), Scaling of peak ground acceleration and peak ground velocity recorded in the Netherlands, *Bollettino di Geofisica Teorica ed Applicata*, *45*(3), 153–168.

Dost, B., E. Ruigrok, and J. Spetzler (2017), Development of probabilistic seismic hazard assessment for the Groningen gas field, *Netherlands Journal of Geosciences*, *96*(5), s235–s245, doi:10.1017/njg.2017.20.

Dost, B., A. van Stiphout, D. Kühn, M. Kortekaas, E. Ruigrok, and S. Heimann (2020), Probabilistic moment tensor inversion for hydrocarbon-induced seismicity in the Groningen gas field, the Netherlands, part 2: application, *Seismological Research Letters*, *submitted*.

Douglas, J., and P. Jousset (2011), Modeling the difference in ground-motion magnitude-scaling in small and large earthquakes, *Seismological Research Letters*, *82*(4), 504–508.

Douglas, J., B. Edwards, V. Convertito, N. Sharma, A. Tramelli, D. Kraaijpoel, B. M. Cabrera, N. Maercklin, and C. Troise (2013), Predicting ground motion from induced earthquakes in geothermal areas, *Bulletin of the Seismological Society of America*, *103*(3), 1875–1897.

Ebel, J., and K.-P. Bonjer (1990), Moment tensor inversion of small earthquakes in south-western Germany for the fault plane solution, *Geophysical Journal International*, *101*(1), 133–146.

Gehl, P., J. Douglas, and D. d'Ayala (2017), Inferring earthquake ground-motion fields with Bayesian Networks, *Bulletin of the Seismological Society of America*, *107*(6), 2792–2808.

Haak, H. (Ed.) (1993), *Summary of the final report on a multidisciplinary study of the relation between Gas production and earthquakes in the northern part of the Netherlands*, p. 16, Royal Netherlands Meteorological Institute (de Bilt), ISBN 90-369-2052-3.

Jonker, A., C. Ostendorf, and C. Cauberg-Huygen (2017), *SBR Trillingsrichtlijn A: Schade aan bouwwerken:2017*, SBRCURnet, https://www.naa.nl/wp-content/uploads/2019/04/SBR-Trillingsrichtlijn-A-Schade-aan-bouwwerken-2017.pdf.

KNMI (1993), Netherlands Seismic and Acoustic Network, Royal Netherlands Meteorological Institute (KNMI), Other/Seismic Network, doi:10.21944/e970fd34-23b9-3411-b366-e4f72877d2c5.

Kraaijpoel, D., and B. Dost (2013), Implications of salt-related propagation and mode conversion effects on the analysis of induced seismicity, *Journal of Seismology*, *17*(1), 95–107.

Kruiver, P. P., et al. (2017), An integrated shear-wave velocity model for the Groningen gas field, the Netherlands, *Bulletin of Earthquake Engineering*, doi: 10.1007/s10518-017-0105-y.

Lienert, B. R., E. Berg, and L. N. Frazer (1986), HYPOCENTER: An earthquake location method using centered, scaled, and adaptively damped least squares, *Bulletin of the Seismological Society of America*, *76*(3), 771–783.

Ntinalexis, M., et al. (2019), Ground-motion networks in the groningen field: usability and consistency of surface recordings, *Journal of Seismology*, *23*, 1233–1253.

Rodriguez-Marek, A., P. P. Kruiver, P. Meijers, J. J. Bommer, B. Dost, J. van Elk, and D. Doornhof (2017), A regional site-response model for the Groningen gas field, *Bulletin of the Seismological Society of America*, *107*(5), 2067–2077.

Ruigrok, E., and B. Dost (2020), Advice on the computation of peak-ground-velocity confidence regions for the Zuidlaren 23-12-2016 event, *KNMI Technical Report*, *TR-382*.

Schiappapietra, E., and J. Douglas (2020), Modelling the spatial correlation of earthquake ground motion: Insights from the literature, data from the 2016–2017 central italy earthquake sequence and ground-motion simulations, *Earth-Science Reviews*, p. 103139.

Spetzler, J., E. Ruigrok, and B. Dost (2018), Improved 3D hypocenter method for induced earthquakes in Groningen, Nederlands Aardwetenschappelijk Congres, March 15-16. Veldhoven, the Netherlands.

Stafford, P. J., A. Rodriguez-Marek, B. Edwards, P. P. Kruiver, and J. J. Bommer (2017), Scenario dependence of linear site-effect factors for short-period response spectral ordinates, *Bulletin of the Seismological Society of America*, *107*(6), 2859–2872.

Stafford, P. J., B. D. Zurek, M. Ntinalexis, and J. J. Bommer (2019), Extensions to the groningen ground-motion model for seismic risk calculations: component-to-component variability and spatial correlation, *Bulletin of Earthquake Engineering*, *17*(8), 4417–4439.

Tarantola, A. (2005), *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia.

Van Ginkel, J., E. Ruigrok, and R. Herber (2019), Assessing soil amplifications in Groningen, the Netherlands, *First Break*, *37*(10), 33–38.

Wald, D. J., B. C. Worden, V. Quitoriano, and K. L. Pankow (2006), Shakemap® manual, *Technical Manual, users guide, and software guide Version.*

Willacy, C., E. van Dedem, S. Minisini, J. Li, J.-W. Blokland, I. Das, and A. Droujinine (2019), Full-waveform event location and moment tensor inversion for induced seismicity, *Geophysics*, *84*(2), KS39–KS57.

Worden, C., D. Wald, T. Allen, K. Lin, D. Garcia, and G. Cua (2010), A revised ground-motion and intensity interpolation scheme for ShakeMap, *Bulletin of the Seismological Society of America*, *100*(6), 3083–3096.

Worden, C. B., E. M. Thompson, J. W. Baker, B. A. Bradley, N. Luco, and D. J. Wald (2018), Spatial and spectral interpolation of ground-motion intensity measure observations, *Bulletin of the Seismological Society of America*, *108*(2), 866–875.

Zhou, H.-w. (1994), Rapid three-dimensional hypocentral determination using a master station method, *Journal of Geophysical Research: Solid Earth*, *99*(B8), 15,439–15,455.